

# **Sensordatenfusion und Bildverarbeitung zur Objekt- und Gefahrenerkennung**

Von der Fakultät für Maschinenbau  
der Technischen Universität Carolo-Wilhelmina zu Braunschweig

zur Erlangung des akademischen Grades

eines Doktor-Ingenieurs (Dr.-Ing.)

genehmigte Dissertation

**von**

Dipl. Inform. Álvaro Catalá Prat

**aus**

Palma de Mallorca (Spanien)

eingereicht am: 18. Juni 2010

mündliche Prüfung am: 19. November 2010

Referenten:

Prof. Dr.-Ing. Karsten Lemmer

Prof. Dr. rer. nat. Ralf Reulke



**Berichte aus dem DLR-Institut für Verkehrssystemtechnik**

**Band 12**

**Sensordatenfusion und Bildverarbeitung zur  
Objekt- und Gefahrenerkennung**

**Álvaro Catalá Prat**

**Herausgeber:**

Deutsches Zentrum für Luft- und Raumfahrt e.V.  
Institut für Verkehrssystemtechnik  
Lilienthalplatz 7, 38108 Braunschweig

**ISSN 1866-721X**

DLR-TS 1.12

Braunschweig, im März 2011

Institutsdirektor:  
Prof. Dr.-Ing. Karsten Lemmer

Verfasser:  
Álvaro Catalá Prat



## **Vorwort des Herausgebers**

Liebe Leserinnen und Leser,

In Ihren Händen halten Sie einen Band unserer Buchreihe „Berichte aus dem DLR-Institut für Verkehrssystemtechnik“. In dieser Reihe veröffentlichen wir spannende, wissenschaftliche Themen aus dem Institut für Verkehrssystemtechnik des Deutschen Zentrums für Luft- und Raumfahrt e.V. (DLR) und aus seinem Umfeld. Einen Teil der Auflage stellen wir Bibliotheken und Fachbibliotheken für ihren Buchbestand zur Verfügung. Herausragende wissenschaftliche Arbeiten und Dissertationen finden hier ebenso Platz wie Projektberichte und Beiträge zu Tagungen in unserem Hause von verschiedenen Referenten aus Wirtschaft, Wissenschaft und Politik.

Mit dieser Veröffentlichungsreihe verfolgen wir das Ziel, einen weiteren Zugang zu wissenschaftlichen Arbeiten und Ergebnissen zu ermöglichen. Wir nutzen die Reihe auch als praktische Nachwuchsförderung durch die Publikation der wissenschaftlichen Ergebnisse von Dissertationen unserer Mitarbeiter und auch externer Doktoranden. Veröffentlichungen sind wichtige Meilensteine auf dem akademischen Berufsweg. Mit der Reihe „Berichte aus dem DLR-Institut für Verkehrssystemtechnik“ erweitern wir das Spektrum der möglichen Publikationen um einen Baustein. Darüber hinaus verstehen wir die Kommunikation unserer Forschungsthemen als Beitrag zur nationalen und internationalen Forschungslandschaft auf den Gebieten Automotive, Bahnsysteme und Verkehrsmanagement.

Die Qualität und Robustheit von Daten zur Objekt- und Gefahrenerkennung als verlässliche Grundlage für Fahrerassistenz zu verbessern, ist das Ziel der im vorliegenden Band veröffentlichten Dissertation. Dazu wurden Daten eines Laserscanners und einer Kamera verarbeitet und fusioniert. Systematisch wurden Algorithmen entwickelt, prototypisch umgesetzt und mit simulierten und realen Daten getestet. Mit diesem Ansatz der Sensordatenfusion und Bilderkennung können Objekte zuverlässig erkannt und verfolgt und atypische Situationen als Gefahren erkannt werden. Auf einem soliden methodischen Fundament basierend ist diese Arbeit ein wichtiger Beitrag für die Entwicklung zuverlässiger Fahrerassistenz – und damit auch für eine sichere Mobilität der Zukunft.

Prof. Dr.-Ing. Karsten Lemmer



## **Vorwort des Autors**

Diese Arbeit entstand während meiner wissenschaftlichen Tätigkeit am Institut für Verkehrssystemtechnik (TS) des DLR in Braunschweig.

Ich möchte mich hiermit bei den vielen Menschen, die mich begleitet und unterstützt haben, bedanken.

Zunächst möchte ich mich recht herzlich bei meinem Doktorvater, Herrn Prof. Dr.-Ing. Karsten Lemmer, für seine Unterstützung und umgängliche Art bedanken. Meinem Zweitprüfer, Herrn Prof. Dr. rer. nat. Ralf Reulke, bin ich ebenso außerordentlich dankbar. Trotz der Entfernung zwischen Braunschweig und Berlin hatten wir einen regen und fruchtbaren Austausch, nicht nur inhaltlich sondern auch was die Begeisterung für eine bestimmte Insel angeht. Als weitere essentielle Rolle möchte ich die von Herrn PD Frank Köster hervorheben. Seine Ankunft im Institut TS bedeutete für mich einen Wendepunkt der Arbeit. Für seine wertvolle Unterstützung, effektive Motivation und stets gute Laune kann ich ihm nur dankbar sein. Ich danke ebenso Herrn Prof. Dr.-Ing. Dr. h.c. mult. Eckehard Schnieder für die freundliche Übernahme des Vorsitzes der Prüfungskommission.

Meinen Kollegen möchte ich hiermit ebenso für die herzliche Aufnahme im Institut TS, für die zahlreichen inhaltlichen Diskussionen sowie für unzählige schöne und lustige Momente danken. Insbesondere weiß ich den regen Austausch mit Sascha Knake-Langhorst zu schätzen. Auch die Arbeit mit Anna Schieben und Firas Lethaus hat mir viel Freude gemacht. Mein weiterer Dank gilt Jan Gačnik, Silke Köhler, Frederik Meysel, Marcus Talg, André Bolles und Hennadiy Netuzhylov. Euch allen, Danke!

Meiner ganzen Familie danke ich auch für ihre herzliche Unterstützung und Motivation aus der Ferne. Durch sie wurden meine Akkus immer wieder aufgeladen, um diese Arbeit Stück für Stück voranzubringen.

Schließlich bedanke ich mich aus tiefstem Herzen bei meinem Freund Arne Burisch. Er hat mich seit dem ersten Tag begleitet, er hat mir zugehört, er war mit mir mal frustriert, mal euphorisch, aber immer bei mir. Für das Inhaltliche, das Logistische und die liebevolle Begleitung: Danke, Gracias. Ich hoffe nur, ich kann dir bald so sehr helfen, wie du mir geholfen hast.

Dipl.-Inform. Álvaro Catalá Prat





# Inhaltsverzeichnis

<b>Vorwort des Herausgebers.....</b>	<b>iii</b>
<b>Vorwort des Autors.....</b>	<b>v</b>
<b>Kurzfassung .....</b>	<b>xvii</b>
<b>Abstract .....</b>	<b>xix</b>
<b>Resumen .....</b>	<b>xxi</b>
<b>1 Einleitung .....</b>	<b>1</b>
1.1 Assistenz und Automation zur Erhöhung der Fahrsicherheit .....	1
1.2 Sensordatenverarbeitung und -fusion als Grundlage für Assistenz und Automation	2
1.3 Der Entwicklungsprozess von Assistenz und Automation.....	4
1.4 Verwendete Sensoren .....	5
1.5 Zielsetzung, Forschungsfragen und Beitrag der Arbeit .....	5
1.6 Gliederung und Überblick .....	8
<b>2 Stand der Technik .....</b>	<b>9</b>
2.1 Umfelderkennung in der Automobilbranche.....	9
2.2 Multi-Target-Tracking .....	10
2.3 Objekterkennung aus einzelnen Sensoren .....	12
2.4 Fusionsansätze zur Objekterkennung .....	17
2.5 Erkennung von Gefahrensituationen .....	22
2.6 Zusammenfassung des Kapitels .....	26
<b>3 Auswahl und Definition des Fusionsansatzes und des Objektmodells .....</b>	<b>29</b>
3.1 Fusionsansatz.....	29
3.2 Modell für Objektbeobachtungen und Objekthypothesen.....	31
3.3 Zusammenfassung des Kapitels .....	34
<b>4 Verarbeitung auf Sensordatenebene .....</b>	<b>35</b>
4.1 Verarbeitung der Kamerabilder .....	35
4.2 Aufbereitung der Laserscannerdaten.....	39
4.3 Korrektur der Fahrzeugschwankungen .....	39
4.4 Ermittlung des Fahrkorridors .....	43

4.5	Zusammenfassung des Kapitels .....	48
<b>5</b>	<b>Objekterkennung aus Bildern.....</b>	<b>49</b>
5.1	Extraktion der Bildstruktur .....	49
5.2	Klassifikation der Bildsegmente .....	51
5.3	Erzeugung von Objektbeobachtungen.....	52
5.4	Validierung und Verfeinerung der Objektbeobachtungen.....	54
5.5	Genauigkeit und Beobachtbarkeit der Objektbeobachtungen.....	56
5.6	Zusammenfassung des Kapitels .....	57
<b>6</b>	<b>Objekterkennung aus Laserscannerdaten .....</b>	<b>59</b>
6.1	Anpassung an ausgerichtetes Quadermodell .....	59
6.2	Beobachtbarkeit der Laserscannerobjekte.....	62
6.3	Genauigkeit der Laserscannerbeobachtungen .....	64
6.4	Zusammenfassung des Kapitels .....	65
<b>7</b>	<b>Objektverfolgung und Fusion .....</b>	<b>67</b>
7.1	Dynamikmodell.....	69
7.2	Prädiktion .....	70
7.3	Variable Referenz und Referenzpunkt .....	71
7.4	Assoziation und Gating.....	73
7.5	Schätzung mittels des Informationsfilters .....	76
7.6	Partielle Beobachtbarkeit der Objekte .....	79
7.7	Zuverlässigkeit einer Objekthypothese .....	86
7.8	Verwaltung und Umgang mit inkonsistenten Objekthypothesen.....	88
7.9	Zusammenfassung des Kapitels .....	92
<b>8</b>	<b>Gefahrenerkennung.....</b>	<b>93</b>
8.1	Statistische Gefahrenerkennung .....	93
8.2	Praktische Umsetzung der Gefahrenerkennung .....	96
8.3	Auslegung der Gefahrenerkennung und Bezug zur Güte der Objektdaten .....	100
8.4	Zusammenfassung und Diskussion der Gefahrenerkennung .....	102
<b>9</b>	<b>Experimentelle Ergebnisse.....</b>	<b>105</b>
9.1	Plattformen, Sensoren und Testfahrten .....	105
9.2	Ergebnisse der Bestimmung der Fahrzeugorientierung .....	106
9.3	Ergebnisse der Fahrkorridorerkennung.....	108

9.4	Ergebnisse der Objekterkennung und -verfolgung .....	110
9.5	Ergebnisse der Gefahrenerkennung .....	124
<b>10</b>	<b>Zusammenfassung und Ausblick.....</b>	<b>127</b>
10.1	Zusammenfassung und Diskussion .....	127
10.2	Ausblick.....	129
<b>Anhang</b>	<b>.....</b>	<b>130</b>
<b>A</b>	<b>Sensormodelle .....</b>	<b>131</b>
A.1	Sensorik und Orientierungsparameter .....	131
A.2	Geometrische Beschreibung einer Kamera.....	133
A.3	Geometrische Beschreibung eines Laserscanners .....	135
<b>B</b>	<b>Geometrische Transformation zwischen der Bild- und der Straßenebene.....</b>	<b>137</b>
<b>C</b>	<b>Der Kalman-Filter und der Informationsfilter .....</b>	<b>139</b>
C.1	Systembeschreibung .....	139
C.2	Der Kalman-Filter .....	140
C.3	Der Informationsfilter .....	143
C.4	Verwendbarkeit des Kalman-Filters für indirekte Beobachtungen.....	147
<b>D</b>	<b>Beispiele der partiellen Beobachtbarkeit .....</b>	<b>151</b>
	<b>Verwendete Abkürzungen und Formelzeichen.....</b>	<b>159</b>
	<b>Literaturverzeichnis .....</b>	<b>165</b>



# Abbildungsverzeichnis

Abbildung 1-1 :	Konzept von Gefährdung und Gefahr am Beispiel von typischen Fahr-situationen.....	2
Abbildung 1-2 :	Großanlagen des Instituts TS: das <i>ViewCar</i> , eine Messbrücke, der dy-namische Fahrsimulator und das <i>FASCar</i> . ....	4
Abbildung 1-3 :	Modulare Prozessierungskette. ....	7
Abbildung 2-1 :	Einige Meilensteine der Sensor- und Erkennungssysteme in der Auto-mobilbranche.....	10
Abbildung 2-2 :	Prozessmodell eines Multi-Target-Tracking-Systems. ....	11
Abbildung 2-3 :	Ansätze zur Objekterkennung aus Laserscannerdaten.....	14
Abbildung 2-4 :	Ansätze zur Objekterkennung aus Mono-Kamerabildern. ....	15
Abbildung 2-5 :	Prinzipien der Fusion.....	17
Abbildung 2-6 :	Abstraktionsebenen der Fusion. ....	19
Abbildung 2-7 :	Beispiele von Fusionssystemen zur Objekterkennung und Verfolgung...	21
Abbildung 2-8 :	Konzept von Gefahr nach unterschiedlichen Theorien.....	23
Abbildung 2-9 :	Ansätze zur Gefahrenerkennung.....	25
Abbildung 3-1 :	Verwendetes ausgerichtetes Quadermodell und mögliche Referenzen...	32
Abbildung 3-2 :	Verändern des Aussehens eines Objektes im Laufe der Zeit und bei unterschiedlichen Anordnungen zu anderen Objekten und zum Sensor.	32
Abbildung 4-1 :	Verarbeitung auf Sensorebene im Rahmen der Arbeit. ....	35
Abbildung 4-2 :	Beispiel für die Kantenfilterung.....	36
Abbildung 4-3 :	Beispiel der Hough-Transformation. ....	37
Abbildung 4-4 :	Effekt von Grauwertstufe $p_{gs}$ und Verschiebungskonstante $p_{go}$ in einem Texturbildauschnitt. ....	38
Abbildung 4-5 :	Texturbild in zwei unterschiedlich beleuchteten Situationen.....	38
Abbildung 4-6 :	Schritte der Bestimmung der Fahrzeugorientierung .....	40
Abbildung 4-7 :	Beispiel der Korrektur der Fahrzeugorientierung.....	42
Abbildung 4-8 :	Beispiele von Fahrsituationen mit unterschiedlichem Fahrkorridor. ....	43
Abbildung 4-9 :	Beitrag der Kameradaten zur Ermittlung des Fahrkorridors.....	45
Abbildung 4-10 :	Beitrag der Karteninformationen zur Ermittlung des Fahrkorridors.....	46
Abbildung 4-11 :	Beitrag der Laserscannerdaten zur Ermittlung des Fahrkorridors.....	47
Abbildung 4-12 :	Beispiele für die Ermittlung der Fahrkorridorbegrenzungen in unter-schiedlichen Fahrsituationen, dargestellt mit dem Tool <i>AroundViewCar</i> .	48
Abbildung 5-1 :	Objekterkennung aus Bildverarbeitung im Rahmen der Arbeit. ....	50
Abbildung 5-2 :	Ein vereinfachtes Texturbild und Ausschnitt des entsprechenden Bild-segmentgraphen. ....	50
Abbildung 5-3 :	Beispiel für die Klassifikation von Bildsegmenten. ....	53
Abbildung 5-4 :	Beispiele für die Validierung von Objektbeobachtungen aus Bildern....	55
Abbildung 5-5 :	Beispiele für die Verfeinerung von Objektbeobachtungen aus Bildern...	55
Abbildung 5-6 :	Räumlicher Vorwärtsschnitt und Genauigkeit der Objektbeobachtungen aus Bildern.....	57

Abbildung 6-1 :	Objekterkennung anhand von Laserscannerdaten im Rahmen der Arbeit.	60
Abbildung 6-2 :	Grundform von Laserscannerobjekten. ....	60
Abbildung 6-3 :	Anpassung von Objektteilen an gerichtete Quader. ....	61
Abbildung 6-4 :	Beispiel für Mehrdeutigkeiten in der Anpassung derselben Laserscanner-Reflexionen auf das gerichtete Quadermodell. ....	62
Abbildung 6-5 :	Bestimmung und Überprüfung der partiellen Beobachtbarkeit von Laserscanner-Objektbeobachtungen. ....	63
Abbildung 6-6 :	Beispiel der Erzeugung von Objektbeobachtungen anhand von Laserscannerdaten. ....	66
Abbildung 7-1 :	Objektverfolgung und Fusion im Rahmen der Arbeit. ....	67
Abbildung 7-2 :	Prozessmodell für das Verfolgungs- und Fusionsmodul zur Übersicht des Kapitels. ....	68
Abbildung 7-3 :	Auswahl der möglichen Referenzen eines Objekts und Veränderung der Genauigkeit nach Referenz. ....	73
Abbildung 7-4 :	Verdeutlichung der euklidischen und der Mahalanobis-Distanz um zwei Zufallsgrößen. ....	75
Abbildung 7-5 :	Objektassoziation und Gating-Strategie. ....	76
Abbildung 7-6 :	Beispiel für Objektverfolgung mit Beobachtbarkeitsbetrachtungen. ....	85
Abbildung 7-7 :	Zustandsautomat einer Objekthypothese. ....	87
Abbildung 7-8 :	Beispiele für Objektverfolgung mit Verdoppelungsstrategie. ....	89
Abbildung 7-9 :	Beispiele für <i>split</i> und <i>merge</i> von Objekten. ....	91
Abbildung 8-1 :	Anwendungsebene im Rahmen der Arbeit. ....	93
Abbildung 8-2 :	Graphische Darstellung des statistischen Informationsgitters (SIG). ....	94
Abbildung 8-3 :	Prozessüberblick für die Gefahrenerkennung. ....	95
Abbildung 8-4 :	Hinzufügung neuer Daten in das SIG gewichtet mit deren Fehlerkovarianz. ....	99
Abbildung 8-5 :	Statistischer Test zum Vergleich der Beobachtungen mit dem Schwellwert im SIG. ....	100
Abbildung 8-6 :	Einfluss der Datengenauigkeit auf die Gefahrenerkennung. ....	101
Abbildung 9-1 :	Validierung der Korrektur von Fahrzeugschwankungen anhand von gesetzten Parametern in der Simulationsumgebung. ....	107
Abbildung 9-2 :	Beispiel der Fahrkorridorerkennung anhand einer Sequenz realer Testdaten. ....	109
Abbildung 9-3 :	Schnappschuss aus dem Experiment TS1. ....	112
Abbildung 9-4 :	Ablauf des Testszenarios TS1 anhand der Kamera, des Laserscanners und der Fusion. ....	113
Abbildung 9-5 :	Schnappschuss aus dem Experiment TS2. ....	114
Abbildung 9-6 :	Ablauf des Testszenarios TS2 anhand der Kamera, des Laserscanners und der Fusion. ....	116
Abbildung 9-7 :	Schnappschuss aus dem Experiment TS3. ....	117
Abbildung 9-8 :	Ablauf des Testszenarios TS3 anhand der Kamera, des Laserscanners und der Fusion. ....	118
Abbildung 9-9 :	Schnappschuss aus dem Experiment RA1. ....	119

Abbildung 9-10 :	Ablauf des Realdatenausschnitts RA1 anhand der Kamera, des Laserscanners und der Fusion.....	120
Abbildung 9-11 :	Schnappschuss aus dem Experiment RA2. ....	122
Abbildung 9-12 :	Ablauf des Realdatenausschnitts RA2 anhand der Kamera, des Laserscanners und der Fusion.....	123
Abbildung 9-13 :	Beispiel einer SIG-Struktur.....	125
Abbildung 9-14 :	Beispiele für die Verteilungsfunktion in zwei unterschiedlichen SIG-Zellen.....	125
Abbildung A-1 :	Kamera und Laserscanner Sensoren mit schematischer Aufbau. ....	132
Abbildung A-2 :	Abbildungsvorgang in einem Sensor unter Berücksichtigung der Alignment-Parameter. ....	132
Abbildung A-3 :	Abbildungsvorgang in einer idealen Lochkamera.....	133





# Tabellenverzeichnis

Tabelle 1-1 : Vergleich der Eigenschaften von Kamera und Laserscanner für die Objekterkennung (siehe Beschreibung im Text).....	6
Tabelle 2-1 : Einschätzung von Sensoren bezüglich unterschiedlicher Komponenten der Objekterkennung.....	16
Tabelle 2-2 : Vergleich der grundsätzlichen Ansätze zur Objekterkennung anhand von Sensorfusion .....	20
Tabelle 5-1 : Beispiel eines Regelsatzes für die Klassifikation eines Bildsegments. ....	52
Tabelle 6-1 : Beispiele zur Entscheidung der Referenz eines Objekts. ....	62
Tabelle 6-2 : Beispiele für die Bestimmung der Beobachtbarkeit der Komponenten einer Objektbeobachtung und entsprechende Änderung der Referenz.....	63
Tabelle 7-1 : Regeln zum Wechseln des Objektszustands <i>st</i> .....	86
Tabelle 9-1 : Quantitative Ergebnisse aus dem Testszenario TS1.....	111
Tabelle 9-2 : Quantitative Ergebnisse aus dem Testszenario TS2.....	114
Tabelle 9-3 : Quantitative Ergebnisse aus dem Testszenario TS3.....	117
Tabelle 9-4 : Quantitative Ergebnisse aus dem Realdatenausschnitt RA1.....	119
Tabelle 9-5 : Quantitative Ergebnisse aus dem Realdatenausschnitt RA2.....	121



## Kurzfassung

Schwerpunkte der vorliegenden Arbeit sind die automatische Erkennung und Verfolgung von Objekten in Fahrsituationen sowie die Ableitung von potentiellen Gefahren. Hierzu werden die Daten eines Laserscanners und einer Kamera verarbeitet und fusioniert. Die Arbeit stellt neue Methoden der unmittelbaren Umfelderkennung dar und dient als Grundlage für innovative Assistenz- und Automationssysteme im Fahrzeug. Solche Systeme unterstützen den Fahrer zur Erhöhung der Fahrsicherheit, der Verkehrseffizienz und des Komforts.

Die entwickelten Methoden sind auf unterschiedlichen Abstraktionsebenen angesiedelt:

Auf Sensordatenebene werden die Daten vorbereitet und reduziert. Insbesondere liegt der Fokus auf der Erkennung von Fahrschwankungen aus Kamerabildern und auf der Erkennung des Fahrkorridors als Interessenbereich aus den Daten mehrerer Sensoren.

Auf Objektebene findet die zentrale Datenfusion statt. Durch die Auswahl einer konkurrierenden Objektfusion wird eine hohe Sensorunabhängigkeit, Zuverlässigkeit und Verfügbarkeit erreicht. Hierzu werden im Vorfeld Objektbeobachtungen aus beiden Sensoren extrahiert. Diese werden zur Objekterkennung und -verfolgung fusioniert, mit besonderem Fokus auf die Robustheit gegenüber manövrierenden Objekten, Messausreißern, *split*- und *merge*-Effekten und der partiellen Beobachtbarkeit der Objekte.

Auf der Anwendungsebene der Arbeit wird die frühzeitige Erkennung von potentiellen Gefahrensituationen vorgestellt. Hierzu wurde ein statistischer Ansatz entwickelt, in dem Gefahren als atypische Situationen behandelt werden. Dieser allgemeingültige und erweiterbare Ansatz wird beispielhaft auf Basis der Objektdaten umgesetzt.

Die vorgestellten Ansätze wurden systematisch entwickelt, prototypisch und modular implementiert sowie mit simulierten und realen Daten getestet. Die Ergebnisse zeigen hierbei eine Steigerung der Qualität und Robustheit der Daten, so dass ein wichtiger Beitrag zur Verbesserung von Assistenz und Automation geleistet wird.



# Abstract

## *Sensor Data Fusion and Image Processing for Object and Hazard Detection*

The present work deals with automatic detection and tracking of objects in driving situations as well as derivation of potential hazards. To do this, the data of a laser scanner and a camera is processed and fused. The work provides new methods in the area of immediate environment detection and modeling. Thus, it creates a basis for innovative driver assistance and automation systems. The aim of such systems is to improve driving safety, traffic efficiency and driving comfort.

The methods introduced in this work can be classified into different abstraction levels:

At sensor data level, the data is prepared and reduced. In this work, the focus is especially set on the detection of driving oscillations from camera images and on the detection of the driving corridor from the data of different sensors, used later as the primary area of interest.

At object level the central data fusion is done. High reliability, availability and sensor independency are achieved by choosing a competitive object fusion approach. As an input of the data fusion, object observations from camera and laser scanner data are extracted. These are then fused at the aim of object detection and tracking, where aspects such as robustness against manoeuvring objects, measurement outliers, split and merge effects, as well as partial object observability are addressed.

At application level, early detection of potential hazards is addressed. A statistical approach has been chosen and developed, in which hazards are handled as atypical situations. This general and expandable approach is exemplarily shown based on the detected object data.

The presented strategies and methods have been developed systematically, implemented in a modular prototype and tested with simulated and real data. The test results of the data fusion system show a win in data quality and robustness, with which an improvement of driver assistance and automation systems can be reached.



## Resumen

### *Fusión de Datos de Sensores y Procesado de Imágenes para la Detección de Objetos y Peligros*

La presente tesis se enfoca en la detección y el seguimiento automático de objetos en situaciones de conducción así como en la derivación de posibles peligros. Para tales efectos, los datos de un escáner láser y de una cámara son procesados y fusionados. El trabajo presenta nuevos métodos de percepción y modelado del entorno inmediato. Éste sirve de fundamento al desarrollo de innovadores sistemas de asistencia y automatización en la conducción. Tales sistemas aspiran a incrementar la seguridad de la conducción, la eficiencia del tráfico y el confort.

Los métodos desarrollados se ubican a diferentes niveles de abstracción:

A nivel de datos de sensores los datos son preprocesados y reducidos. El foco está en el reconocimiento de las oscilaciones de la conducción en base a las imágenes de la cámara así como en el reconocimiento del corredor de conducción actual en base a una fusión de los sensores, usado más adelante como area de interés.

A nivel de objetos se lleva a cabo la fusión central de datos. Gracias a un planteamiento competitivo de fusión de objetos es posible lograr una alta independencia de los sensores, fiabilidad y disponibilidad. Para ello se extraen primero observaciones de objetos del escáner láser y la cámara. Éstas son fusionadas con el fin del reconocimiento y seguimiento de objetos, con especial interés en la robustez frente a objetos maniobrantes, a valores dudosos, a efectos de partido y mezclado de objetos y a la observabilidad parcial de los objetos.

A nivel de aplicación la pronta detección de posibles peligros es tratada. Para ello, un enfoque estadístico ha sido desarrollado en el cual los peligros son tratados como situaciones atípicas. Esta general y ampliable estrategia es ejemplificada con los datos de la detección de objetos.

Los métodos presentados han sido desarrollados sistemáticamente, implementados prototípicamente y testados con datos simulados y reales. Los resultados de las pruebas muestran un incremento de la calidad y robustez de los datos, con lo se puede lograr una mejora en el desarrollo de sistemas de asistencia y automatización en la conducción.





# 1 Einleitung

Assistenz- und Automationssysteme in Fahrzeugen dienen insbesondere der Erhöhung der Fahr-sicherheit und der Verkehrseffizienz sowie der Verbesserung des Fahrkomforts. Für diese Funktio-nalitäten werden Informationen über das Umfeld, das Fahrzeug und den Fahrer benötigt. In dieser Arbeit wird der Fokus auf die Umfelderkennung gelegt, wobei die Sensordatenverarbeitung und -fusion zentrale Bausteine darstellen. Hierbei lehnt sich die Arbeit an den am Institut für Ver-kehrssystemtechnik (*Transportation Systems*, TS) des DLR<sup>1</sup> etablierten Entwicklungsprozess für Assistenz- und Automationssysteme an. Dieser beinhaltet die Analyse von Fahrsituationen, die Erprobung von Assistenz und Automation in simulierter Umgebung sowie deren Umsetzung und Erprobung in realen Systemen.

## 1.1 Assistenz und Automation zur Erhöhung der Fahrsicherheit

Individuelle Mobilität ist zweifelsohne ein wichtiger Bestandteil des Lebens. Durch das stetig anwachsende Verkehrsaufkommen werden oft die Sicherheit, die Effizienz und der Komfort im Kontext des Verkehrs negativ beeinflusst. Durch Assistenz und Automationssysteme wird die-sen Effekten entgegengewirkt, wobei die Erhöhung der Fahrsicherheit an erster Stelle steht. Unterschiedliche Fahrsituationen, wie die in Abbildung 1-1 schematisch zusammengefasst, beinhalten Gefährdungen, die gegebenenfalls zu echten Gefahren bis hin zu Unfällen führen kön-nen. Beispiele sind: eine Kollision mit einem Fahrzeug im toten Winkel (1), ein Auffahrunfall (2), das unbewusste Verlassen der Fahrbahn (3), der Kontrollverlust bei Glatteis (4) oder über-höhter Geschwindigkeit (5), die Vorfahrtmissachtung an einer Kreuzung (6) oder die Kollision mit einem Fußgänger (7).

Eine Gefahr wird anhand ihrer Eintrittswahrscheinlichkeit und ihrer Folgen (z.B. eines Schadens) bewertet. Eine solche Bewertung hängt von zahlreichen Aspekten ab, wie beispielsweise der Dynamik der Szene, den Fahrern und deren Zustand, sowie der Infrastruktur.

Durch auf diese Aspekte bzw. einzelne Faktoren ausgerichtete Assistenz und Automation in Fahrzeugen können Unfälle vermieden oder wenigstens deren Folgen verringert werden, was jeweils der Reduktion ihrer Eintrittswahrscheinlichkeit bzw. ihres Schadenumfangs entspricht.

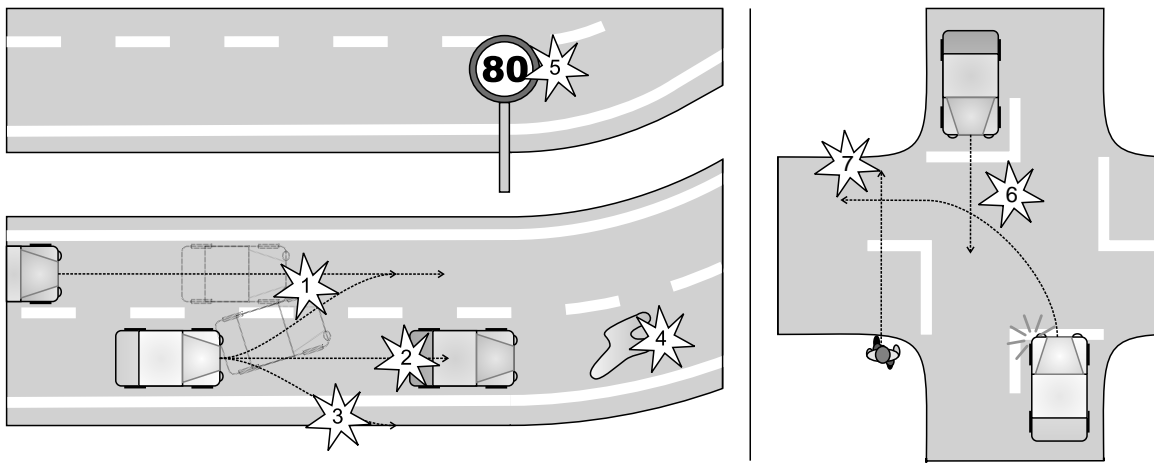
Es gibt zahlreiche Beispiele für solche Systeme, die in Serienfahrzeugen vermehrt zu finden sind. Passive Sicherheitssysteme wie Gurtstraffer oder Airbag sind inzwischen weitgehend Standard. Für viele aktive Systeme gilt dies ebenso, wobei gerade in diesem Segment aktuell ein rasches Wachstum zu beobachten ist. Beispiele für aktive Systeme sind Antiblockiersystem (ABS) und Elektronisches Stabilitätsprogramm (ESP) zum Stabilisieren des Fahrzeugs, sowie Abstandsre-gelautomat (ACC<sup>2</sup>), Spurhalteassistent (LDW<sup>3</sup>) und Notbremsassistent (NBA) zum Führen des Fahrzeugs.

---

<sup>1</sup>DLR - Deutsches Zentrum für Luft- und Raumfahrt e.V. in der Helmholtz Gemeinschaft.

<sup>2</sup>ACC – *Adaptive Cruise Control*

<sup>3</sup>LDW – *Lane Departure Warning*



**Abbildung 1-1:** Konzept von Gefährdung und Gefahr am Beispiel von typischen Fahrsituationen (siehe Beschreibung im Text).

Um ihre Funktion zu erfüllen, verfügen Assistenz- und Automationssysteme im Allgemeinen über Sensorsysteme zur Gewinnung von Informationen über die Fahrsituation. Diese Informationen betreffen das Umfeld, das Fahrzeug und den Fahrer. Beispielsweise erkennt ein ACC ein führendes Fahrzeug anhand eines Radars oder Lidars. Immer häufiger ist der Einsatz von Kameras zur Umfelderkennung zu beobachten, da es sich bei diesen um passive und preiswerte Sensoren handelt. Mit einer Kamera lassen sich unter anderen Informationen zu Fahrstreifenmarkierungen, umgebenden Objekten oder Verkehrszeichen erkennen. Zusätzlich können ebenso Informationen aus digitalen Karten oder aus der Kommunikation mit anderen Verkehrsteilnehmern oder mit der Infrastruktur verwendet werden (*Car2Car* oder *Car2Infrastructure*). Ein ESP benötigt Beschleunigungs- und Drehzahlsensoren zur Abschätzung des Fahrzeugzustands. Müdigkeitserkennungssysteme verwenden NIR-Kameras<sup>4</sup>, um das Gesicht des Fahrers zu erfassen.

Die steigende Zahl an komplexen Assistenz- und Automationssystemen erfordert immer leistungsfähigere Erkennungssysteme, die ein komplettes Bild der Fahrsituation im Umfeld des Fahrers mit hoher Zuverlässigkeit, Genauigkeit und Verfügbarkeit liefern. Zusätzlich bestehen die Herausforderungen darin, die Anzahl von Erkennungssystemen und deren Komponenten in Grenzen zu halten, diese zu vernetzen und ein einheitliches und dabei insbesondere widerspruchsfreies Bild der Fahrsituation an alle Anwendungen in einem Fahrzeug zu liefern.

## 1.2 Sensordatenverarbeitung und -fusion als Grundlage für Assistenz und Automation

Die Bestimmung der Informationen über die Fahrsituation geschieht im Kontext von Assistenz- und Automationssystemen in Fahrzeugen mit geeigneten Sensoren und Methoden. Zu jeder Information können die Daten von Einzelsensoren verarbeitet oder mehrere Sensoren in bestimmter Kombination gemeinsam betrachtet werden, was unter der Bezeichnung Sensordatenfusion zusammengefasst wird. Durch diesen Ansatz werden im Allgemeinen folgende Vorteile erzielt:

- Erweiterung der erfassbaren Informationen: Die Informationen zweier Sensoren können zusammengefügt werden, um neue Dimensionen, Merkmale, Eigenschaften und Arten von

<sup>4</sup>NIR-Kamera: Nahinfrarot-Kamera

Gefahren zu erkennen, die als Grundlage für Automations- und Assistenzsysteme dienen.

- Erweiterung des Erfassungsbereichs: Der Erfassungsbereich zweier oder mehr Sensoren wird zusammengefügt, um eine größere Reichweite oder Öffnungswinkel zu erreichen.
- Erhöhung der Genauigkeit: Durch Beobachtungen einer Größe durch mehrere Sensoren kann die Genauigkeit erhöht (bspw. Positionsgenauigkeit oder Herkunftsrichtung einer Gefährdung) und damit eine bessere Assistenz oder Automation erreicht werden.
- Erhöhung der Zuverlässigkeit: Mehrere Sensoren führen zu einer höheren Zuverlässigkeit der Daten. Die Raten von *false positives* und *false negatives* können reduziert werden.<sup>5</sup>
- Erhöhung der Verfügbarkeit: Die Folgen von temporären Einschränkungen der Aufnahmesysteme (von niedrigen Datenraten bis hin zu kurzzeitigen Ausfällen) können durch Sensordatenfusion reduziert werden.
- Erhöhung der Robustheit: Unter Robustheit eines Systems versteht sich dessen Fähigkeit, seine Funktion auch bei Schwankungen der Umgebungsbedingungen aufrecht zu erhalten (bspw. Robustheit gegen Ausreißer, Ausfall eines Sensors und Rauschen in den Daten).

Es ist ersichtlich, dass durch die Verwendung von Sensoren mit unterschiedlichen physikalischen Erfassungsprinzipien diese Vorteile besonders zum Tragen kommen können.

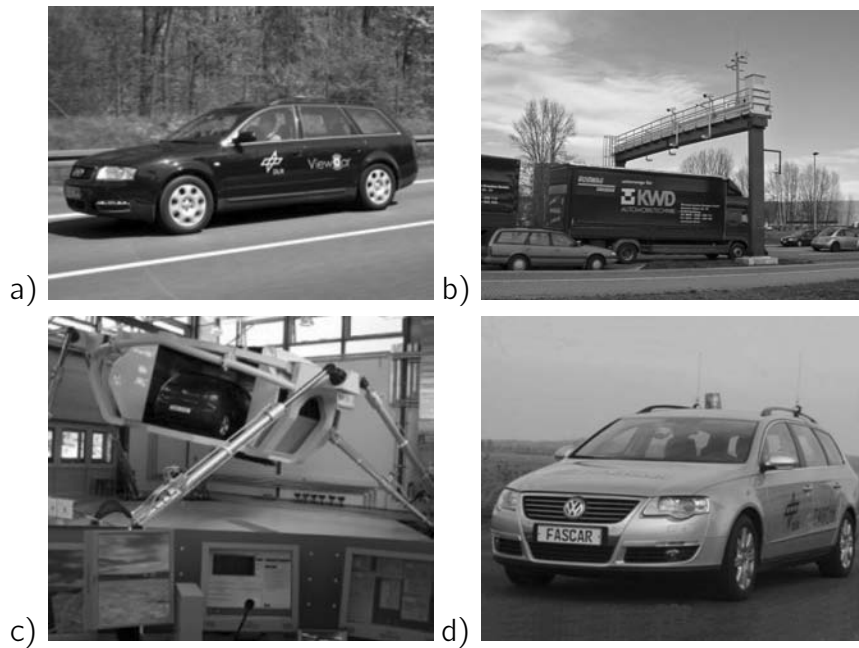
Nichtsdestotrotz treten bei der Sensordatenfusion in der Praxis einige Probleme auf, die zu berücksichtigen sind. Die Fusion mehrerer Sensoren geht stets mit einer Erhöhung der Komplexität des Systems einher. Dadurch kann die Zuverlässigkeit der Ausgabe beeinträchtigt werden. Des Weiteren sind sowohl die Datenrate als auch die benötigte Rechenzeit in einem Fusionssystem höher als bei einzelnen Sensoren, was zu einer Reduktion der Verfügbarkeit führen kann. Die Verwendung mehrerer Sensoren erfordert hochgenaue räumliche und zeitliche Referenzierungsinformationen. Werden diese im Laufe der Zeit ungenauer, so wird die Genauigkeit und Zuverlässigkeit des Fusionssystems ebenso gefährdet.

Im Zusammenhang mit der in dieser Arbeit verfolgten Objekterkennung finden sich die hier diskutierten Ziele und Schwierigkeiten wieder. Zum Beispiel wird die detaillierte Erkennung aller Objektarten (mit Klassifikation und Eigenschaften) angestrebt. Die Position der Objekte soll dabei sehr genau geschätzt und mit einer hohen Frequenz verfolgt werden. Die Dynamik der Objekte lässt sich dabei nicht immer mit einfachen Modellen beschreiben. Da in der Regel *false positives* auftauchen können, ist es von großem Wert, dass diese keinen negativen Einfluss auf die Verfolgung der tatsächlich vorhandenen Objekte haben.

Analog dazu werden bei der Gefahrenerkennung diese Ziele und Schwierigkeiten sichtbar. Beispielsweise können durch Fusion Gefahren frühzeitiger festgestellt werden, als bei einzelnen Sensoren. Durch die zuverlässigeren Daten aus einer Fusion kann die Aufgabe der Gefahrenerkennung ebenso zuverlässiger durchgeführt werden. Allerdings muss auch mit zum Teil widersprüchlichen Daten gerechnet werden, die in der Fusion geklärt werden sollen.

---

<sup>5</sup>Ein *false negative* ist ein Sensorfehler (genannt Typ I), bei dem ein existierendes Objekt durch den Sensor nicht erfasst wird. Ein *false positive* ist ein Sensorfehler (genannt Typ II), bei dem der Sensor Werte liefert, obwohl sich kein Objekt in dem Erfassungsbereich befindet (Phantomobjekte). Diese Fehler können beispielsweise durch Sensorverschmutzung, die physikalischen Objekteigenschaften oder das Signal-Rausch-Verhalten (*signal-to-noise ratio*, SNR) hervorgerufen werden.



**Abbildung 1-2:** Großanlagen des Instituts TS: a) *ViewCar*, b) Messbrücke, c) dynamischer Fahrsimulator und d) *FASCar*.

### 1.3 Der Entwicklungsprozess von Assistenz und Automation

Am Institut TS umfasst ein üblicher Entwicklungsprozess für Assistenz- und Automationssysteme mehrere Schritte, von der Analyse der Fahrsituationen im realen Verkehr, über die Erprobung in simulierter Umgebung, bis hin zur Umsetzung in realer Welt bzw. realen Fahrversuchen.

Ein typischer Design- und Entwurfsprozess beginnt mit der Identifikation des Bedarfs für Assistenz und Automation. Hierfür werden Unfälle sowie das Fahrerverhalten in spezifischen Situationen analysiert. Für diesen Zweck verfügt das Institut TS über unterschiedliche Versuchsplattformen und infrastrukturelle Messeinrichtungen – siehe Beispiele in Abbildung 1-2. Beispielsweise werden mit dem in dieser Arbeit verwendeten Versuchsfahrzeug *ViewCar* (Abbildung 1-2.a) Testfahrten im realen Verkehr durchgeführt. Das Fahrzeug verfügt über geeignete Sensoren zur Erfassung von Informationen über das Umfeld, das Fahrzeug und den Fahrer. Im Zusammenhang mit diesem Prozessschritt werden die Methoden zur automatischen Erkennung von Fahrsituationen genutzt und kontinuierlich weiterentwickelt sowie mithilfe geeigneter Referenzsysteme geprüft.

Auf Basis dieser Informationen entstehen neue Konzepte für Assistenz und Automation. Diese werden zunächst in der virtuellen Welt modelliert und implementiert. Hierfür stehen am Institut TS mehrere Fahrsimulatoren zur Verfügung (siehe exemplarisch Abbildung 1-2.c). Hier erfolgen eine erste Erprobung sowie die Evaluation aller maßgeblichen Aspekte in einer kontrollierten Versuchsumgebung. Im Simulator können bestimmte Fahrsituationen aus Sicherheits-, Wiederholbarkeits- oder Aufwandsgründen umfassend und sinnvoller getestet werden, als im realen Verkehr. Auch die Methoden der Fahrsituationserkennung werden in der Simulation eingesetzt, validiert und getestet.

Ein letzter Schritt in der Prozesskette besteht in der prototypischen Umsetzung gereifter Kon-

zepte in der realen Welt, um sie dort zu validieren. Für diesen Zweck wurde beispielsweise das *FASCar* aufgebaut (siehe Abbildung 1-2.d). Ebenso wie die Sensorik und die Erkennungssysteme findet sich hier auch eine Aktuatorik für aktive Eingriffe in die Fahrdynamik.

Um den Austausch von Daten und Informationen und die Übertragbarkeit der Methoden zwischen den Versuchsanlagen zu ermöglichen, wurde im Institut TS eine einheitliche dienstbasierte Systemarchitektur entwickelt [61]. Hiermit können die Informationen in Diensten verarbeitet und zur Verfügung gestellt werden (wie bspw. unterschiedliche Umfeld- oder Gefahrenerkennungsmethoden). Dabei sind die entwickelten Systeme von Sensoren, Plattformen und Konfigurationen unabhängig und auf unterschiedlichen Versuchsträgern übertragbar. Ein weiterer Vorteil des Einsatzes dieser Systemarchitektur ist die Festlegung eines einheitlichen Modells der Fahrsituation. Die Sensordatenverarbeitung bzw. -fusion kann somit an einer einzigen Stelle durchgeführt und die Ergebnisse von mehreren Anwendungen verwendet werden. Damit wird das Einfügen oder Austauschen von Systemen erleichtert.

Die vorliegende Arbeit wird an den hier beschriebenen Entwicklungsprozess und die Systemarchitektur angelehnt. Hierzu wird das *ViewCar* für die Problemanalyse und prototypische Implementierung der entwickelten Methoden für Objekt- und Gefahrenerkennung verwendet. Die Simulationsumgebung wird für genaue Untersuchungen der Methoden in kontrollierten und wiederholbaren Bedingungen eingesetzt. Insbesondere kann in der Simulationsumgebung die Gefahrenerkennung umfassend untersucht werden. Hierbei können Gefahren durch die Verkehrssimulation gezielt erzeugt und der Ansatz der Gefahrenerkennung validiert werden. In weiteren Schritten können die Methoden der Objekt- und Gefahrenerkennung an ein reales Fahrzeug (beispielsweise das *FASCar*) angepasst und validiert werden, wie im Ausblick dieser Arbeit erwähnt wird.

## 1.4 Verwendete Sensoren

In dieser Arbeit werden ein Laserscanner und eine Mono-Kamera ausgerichtet auf den vorderen Fahrzeugbereich verwendet. Der Laserscanner bietet eine hohe Zuverlässigkeit, Entfernungsgenauigkeit und Reichweite, liefert jedoch außer der Entfernung und Reflektanz einzelner Punkte keine weiteren direkten Informationen über die Objekte. Zusätzliche Informationen, wie zum Beispiel Geschwindigkeit und Größe, werden durch eine entsprechende Verarbeitung der Rohdaten gewonnen. Die Kamera zeichnet mehr Information über die Szenerie auf (Formen, Linien, Texturen, Schatten usw.) und hat zudem in der Regel eine höhere Winkelauflösung als der Laserscanner. Die Informationen müssen jedoch über eine Bildverarbeitung gewonnen werden, von der letztendlich die Erkennungswahrscheinlichkeit abhängt. Zudem reduziert sich die Erkennungsgenauigkeit mit zunehmender Entfernung.

Tabelle 1-1 listet die Stärken und Schwächen beider Sensoren auf und verdeutlicht dabei, wie sie sich durch Fusion gegenseitig ergänzen können.

## 1.5 Zielsetzung, Forschungsfragen und Beitrag der Arbeit

Das Ziel dieser Arbeit besteht in der automatischen Erkennung und Verfolgung von dynamischen und statischen Objekten im unmittelbaren Umfeld des Fahrzeugs sowie der Ableitung potentieller

**Tabelle 1-1:** Vergleich der Eigenschaften von Kamera und Laserscanner für die Objekterkennung (siehe Beschreibung im Text).

Eigenschaft	Kamera	Laserscanner
Entfernungs- genauigkeit	-	++
Winkelauflösung	++ (wenn sichtbare Kanten)	+
Erkennung der Objektgröße	Breite×Länge×Höhe (je nach Blickwinkel)	Breite×Länge (je nach Blickwinkel)
Objekt- klassifikation	++ (nach Bildmerkmalen)	- (nach Größe und Bewegungsmuster)
Zuverlässigkeit ( <i>false positives</i> )	o (je nach Methode)	+
Zuverlässigkeit ( <i>false negatives</i> )	o (je nach Methode)	+
Verfügbarkeit	o (hohe Framerate aber hoher Rechenaufwand)	+
Abhängigkeit von Lichtverhältnissen	-- (Gegenlicht, Brücken oder Nacht)	++
Abhängigkeit von Wetterverhältnissen	-- (Regen und Nebel)	- (Dichter Nebel)

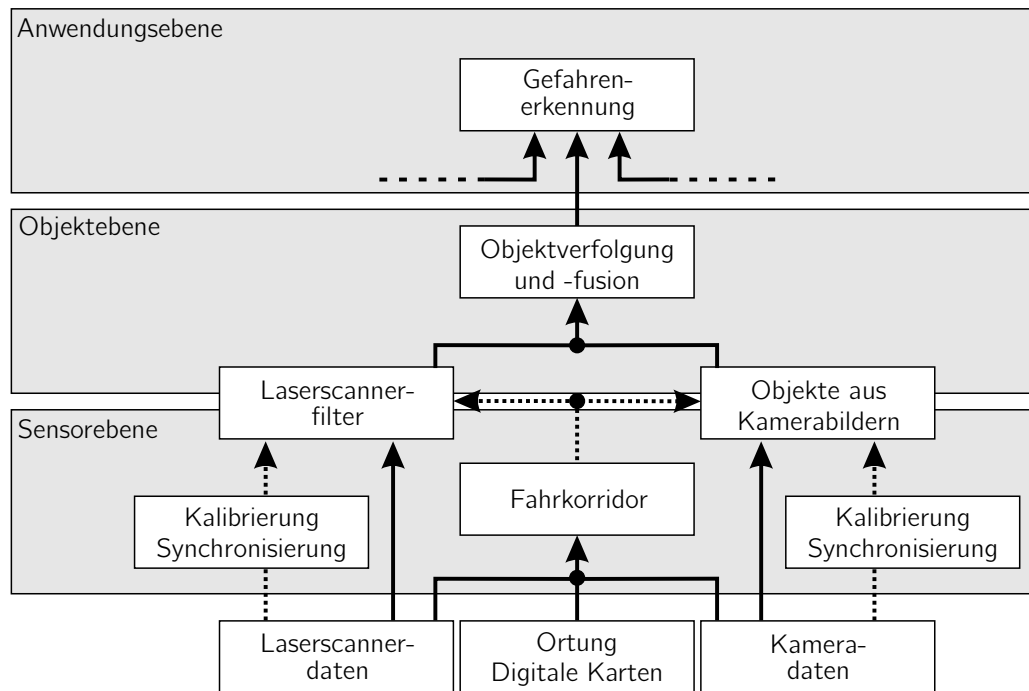
Gefahrensituationen. Die Objekterkennung wird anhand der Verarbeitung und Fusion der Daten einer Kamera und eines Laserscanners erreicht. Ortungsdaten und Informationen aus einer digitalen Karte sind ebenfalls vorhanden und werden unterstützend verwendet. Die darauf aufbauende Gefahrenerkennung erfolgt in einer abstrakten Ebene und wird universell gehandhabt. Sie wird exemplarisch anhand von Objektdaten vorgestellt und dient damit als Eignungskriterium der entwickelten Objekterkennungsmethoden.

Die zentrale Fragestellung der Arbeit ist folgende: welche Chancen und Probleme stellt die Fusion der Daten einer Kamera und eines Laserscanners in der Detektion von Objekten im unmittelbaren Umfeld des Fahrzeugs im Blick auf die frühzeitige Erkennung von potentiellen Gefahren in der Fahrsituation dar?

Diese Frage wird in folgende weitere technische Fragen aufgeschlüsselt: Wie können die Daten aus einem oder unterschiedlichen Sensoren für die Objekterkennung und Verfolgung so verarbeitet bzw. fusioniert werden, dass qualitativ hochwertige Objekthypothesen entstehen? Diese sollen zudem robust sein, gegenüber

- sich verändernden Perspektiven der Objekte,
- manövrierenden Objekten
- und unsicheren aktuellen Sensordaten.

Wie kann die notwendige Verarbeitung der Kamera- und Laserscannerdaten die Anforderungen der Objekterkennung weitestgehend erfüllen? Wie kann die Menge der zu verarbeitenden Sensordaten auf den wesentlichen Bereich reduziert werden?



**Abbildung 1-3:** Modulare Prozessierungskette.

Wie müsste die Architektur einer allgemeingültigen aber gleichzeitig differenzierten Gefahrenerkennung aufgebaut sein, die möglichst frühzeitig auf entsprechende Ereignisse reagiert? Wie hängt die Güte der Objektdaten mit der der Gefahrenerkennung zusammen?

In der Beantwortung dieser eng vernetzten Menge von Fragen liegt der wissenschaftliche Beitrag der vorliegenden Arbeit. In diesem Zusammenhang wird die vollständige Prozessierungskette von den Sensordaten, über die Objekthypothesen bis hin zur Gefahrenerkennung vorgestellt. Diese Prozessierungskette ist in Form eines modularen Aufbaus in Abbildung 1-3 dargestellt. Die Methoden sind auf unterschiedlichen Abstraktionsebenen angesiedelt, nämlich der Sensordatebene, der Objektebene und der Anwendungsebene.

Den Kern der Arbeit bildet die Objektebene mit der Verfolgung und der Fusion der Sensordaten. Durch die Auswahl einer konkurrierenden Fusion auf Objektebene wird eine hohe Sensorunabhängigkeit, Zuverlässigkeit und Verfügbarkeit erreicht. Hierzu ist es notwendig, Objektbeobachtungen aus den Kamera- und den Laserscannerdaten zu extrahieren. Zur Objektextraktion und -verfolgung werden ein einfaches Objektmodell und ein Satz von neuartigen, aufeinander abgestimmten Strategien und Methoden entwickelt, die Lösungen zu den oben aufgelisteten Problemstellungen darstellen. Dazu zählen beispielsweise die Definition einer variablen, logischen Referenz für die Objekthypothesen, die Betrachtung ihrer partiellen Beobachtbarkeit, die Adaption des Mess- und Prozessrauschens, die mehrfache Assoziation zu Beobachtungen und die Verdoppelung bzw. Vereinigung von Objekthypothesen. Damit wird durch die Fusion trotz erhöhter Komplexität ein Gewinn gegenüber herkömmlichen Ansätzen erreicht.

Die Objektextraktion aus den Sensordaten wird auf diese Strategien abgestimmt. Einerseits wird eine auf Texturmerkmalen basierte Bildverarbeitungsmethode vorgestellt und entwickelt. Diese neuartige Methode beruht auf der Annahme, dass sich Objekte vom Hintergrund abheben. Somit liefert sie selbst bei qualitativ beschränkten Kamerabildern und bei verändernder Beleuchtung

zuverlässige Objektbeobachtungen. Die Laserscannerdaten werden anhand einer neuartigen Methode basierend auf dem inneren Winkel der Konturpunkte an das Objektmodell angepasst. Der Vorteil dieser schlanken Methode ist, dass Objektformen differenzierter als bei herkömmlichen Ansätzen klassifiziert und Objekte kompakt dargestellt werden können.

Auf Sensorebene werden in dieser Arbeit zwei Methoden von besonderem Wert entwickelt. Einerseits werden Fahrzeugschwankungen beim Fahren anhand einer neuen Bildverarbeitungsmethode erkannt und korrigiert. Diese basiert auf den bekannten geometrischen Eigenschaften der Fahrstreifenmarkierungen. Andererseits wird der Fahrkorridor anhand einer neuartigen sensornahen, kooperativen Fusion bestimmt, mit dem die Datenmenge in der Objekterkennung auf den wesentlichen Bereich reduziert werden kann.

Schließlich wird auf Anwendungsebene ein Ansatz zur Gefahrenerkennung vorgestellt. Dieser basiert auf der Annahme, dass Gefahren aus atypischen Ereignissen abgeleitet werden können, welche *a priori* eine geringe Eintrittswahrscheinlichkeit aufweisen. Dafür wird ein Fusionskonzept vorgestellt, das statistische Informationen über den normalen Zustand der Daten (bspw. Objektdaten) in entsprechenden Strukturen enthält. Dabei wird die Datensicherheit, Genauigkeit und Beobachtbarkeit, bezogen auf den betrachteten Sensor, berücksichtigt. Die statistischen Strukturen ermöglichen, dass universell gültige Regeln verwendet werden können. Durch deren differenzierte Auslegung können Kriterien, die eine Warnung auslösen (Schwellwerte), adaptiv gestaltet und damit Gefahren frühzeitig erkannt werden. Dieser Ansatz stellt eine vielversprechende Alternative zu herkömmlichen auf der Ursache-Wirkungs-beziehung basierten Systemen dar.

## 1.6 Gliederung und Überblick

Die Arbeit ist wie folgt gegliedert. In diesem einleitenden Kapitel wurde die Thematik der Assistenz und Automation eingeführt, wobei auf die Informationen der Fahrsituation und die Gefahren fokussiert wurde. Dabei wurde die Rolle der Erfassung der Fahrsituation und der Sensordatenfusion hervorgehoben. Anschließend wurden die Ziele und Forschungsfragen dieser Dissertation beschrieben.

Im zweiten Kapitel ist eine Beschreibung zum Stand der Technik im Bereich der Objekt- und Gefahrenerkennung in Bezug auf die Fahrsituation zu finden.

Das dritte Kapitel stellt die Grundentscheidung zur Fusionsansatz dar und definiert das Objektmodell, worauf die restliche Arbeit basiert. Kapitel 4 beschäftigt sich mit der Verarbeitung der Sensordaten auf Sensorebene. Kapitel 5 und 6 gehen auf die Objekterkennung jeweils aus Kamera- und Laserscannerdaten ein. Darauf aufbauend wird im Kapitel 7 die Verfolgung und Fusion der erkannten Objektbeobachtungen in Objekthypothesen erläutert. Im Kapitel 8 ist die Beschreibung des Ansatzes zur Gefahrenerkennung zu finden.

Nachdem im Kapitel 9 die durchgeführten Experimente und deren Ergebnisse dargestellt sind, schließt Kapitel 10 die Arbeit mit einer Zusammenfassung und einem Ausblick ab.

In Ergänzung zu den erwähnten Kapiteln sind einige Anhänge beigefügt, die dem Gesamtverständnis der Arbeit dienen. Ein Bezug darauf erfolgt an den entsprechenden Stellen im Text.



## 2 Stand der Technik

In diesem Kapitel wird der Stand der Technik im Bereich der Umfeld-, der Objekt- und der Gefahrenerkennung erläutert. Für die Objekterkennung wird zuerst in Abschnitt 2.2 das allgemeine Multi-Target-Tracking-Schema eingeführt, das als Referenz für die Recherche dient. Die Objekterkennung wird in einzelsensorbasierte (Abschnitt 2.3) und fusionsbasierte Systeme (Abschnitt 2.4) unterteilt. Die Einzelsensorsysteme werden nach dem jeweils verwendeten Sensor und die Fusionsansätze nach der Abstraktionsebene gegliedert. Im weiteren Verlauf werden Ansätze vorgestellt, die aus den detektierten Daten der Umgebung abstrakte Informationen über mögliche Gefahrensituationen ableiten (Abschnitt 2.5).

### 2.1 Umfelderkennung in der Automobilbranche

Technische Systeme werden immer häufiger mit der Fähigkeit zur Umgebungswahrnehmung ausgestattet. Dabei liegt der Fokus auf Informationen zur Ortung und Lage des Systems, zum Untergrund, zu den umgebenden Objekten und deren Verhalten sowie zur Bedeutung dieser Informationen für die Systemsicherheit. Es können dabei einzelne oder mehrere Sensoren in einem Fusionssystem verwendet werden.

Im militärischen Bereich wurden die ersten modernen Sensorfusionsaufgaben bearbeitet. Die Anwendungen umfassen Ortung von Militärfahrzeugen, Zielbestimmung, Angriffserkennung und Trajektorienplanung. Anwendungsbeispiele sind in der Literatur zu finden, wie zum Beispiel bei Hall [70] und Klein [85]. In der Robotik sind analoge Forschungsfragen vor allem in den Anwendungsbereichen der autonomen mobilen Roboter und der Fernerkundung zu finden.

Die Forschung in der Automobilindustrie zeichnet sich durch zwei Trends im Bereich der Umfelderkennung aus. Der Eine setzt sich immer mehr für die Fusion der einzelnen Sensorinformationen ein. Der Andere zielt auf immer zuverlässigere Daten eines Einzelsensorsystems ab. Abbildung 2-1 veranschaulicht diese Trends anhand einiger wichtiger Meilensteine der letzten Jahre. In der Aufstellung sind viele EU-Projekte und Rahmenprogramme vertreten. Sie differenzieren sich nach den jeweiligen übergeordneten Zielen. So beschäftigte sich das *Eureka-Prometheus*-Projekt (1987-1995) mit dem Ziel des autonomen Fahrens, während sich das Projekt *Chameleon* [58] (2000-2002) auf Pre-Crash-Situationen fokussierte. Das Projekt *Carsense* [68, 90] (2001-2003) wiederum konzentrierte sich auf Umfelderkennung in komplexen Situationen. Das Projekt *Safe-U* [103] (2002-2005), die Initiative *eSafety* [3] (2002-2010) sowie das integrierte Projekt *PREVENT* [2, 129] (2003-2008) haben ihre Ziele hauptsächlich im Bereich der Sicherheit. Auch zu nennen sind vom BMBF<sup>1</sup> bzw. BMWi<sup>2</sup> finanzierte Projekte, wie die Forschungsinitiative *INVENT* [1] (2001-2005), fokussiert auf Fahrerassistenz, aktive Sicherheit und Verkehrsmanagement.

Viele Forschungseinrichtungen haben ebenso Beiträge zur Umfelderkennung geleistet. Dazu zählen Universitäten wie beispielsweise die Stanford University mit dem Fahrzeug *Stanley* (zusammen mit der Volkswagen AG. Abbildung 2-1.a, Thrun et al. [150]), die Technische Universität

---

<sup>1</sup>BMBF: Bundesministerium für Bildung und Forschung

<sup>2</sup>BMWi: Bundesministerium für Wirtschaft und Technologie



a) Stanley



b) VaMoRs



c) VAMP



d) ARGO

**Abbildung 2-1:** Einige Meilensteine der Sensor- und Erkennungssysteme in der Automobilbranche. Siehe Bezug im Text.

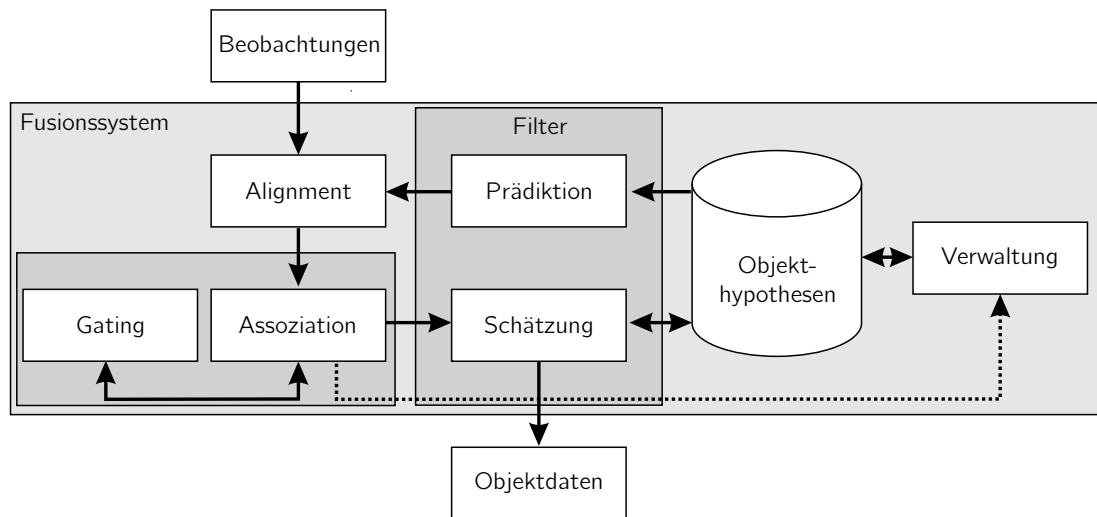
Braunschweig mit dem Versuchsfahrzeug *Caroline* (Form et al. [52, 15]), die Bundeswehr Universität München in Zusammenarbeit mit Mercedes-Benz (Dickmanns [40]) unter anderen mit den Versuchsfahrzeugen *VaMoRs* (Abbildung 2-1.b), *VAMP* (Abbildung 2-1.c) und *VITA-II*, die Universität Ulm (Dietmayer et al. [41]), die Universität Karlsruhe mit dem Fahrzeug *AnnieWAY* (Stiller et al. [139, 142, 141]), die Universitäten von Pavia und Parma mit dem *Centro Ricerche Fiat* (Bertozzi, Broggi et al. [14]) mit *ARGO* (Abbildung 2-1.d). Auch Automobilhersteller und Zulieferer sind hier zu nennen, wie die bereits erwähnten sowie BMW [164], MAN [4], DaimlerChrysler [39] und Ibeo [58]. Die Beiträge dieser Projekte und Institutionen zur Objekt- und Gefahrenerkennung werden ausführlich in den nächsten Abschnitten erläutert.

## 2.2 Multi-Target-Tracking

In dieser Arbeit liegt der Fokus in der Erkennung und Verfolgung von Objekten in Fahrsituationen. Diese Aufgabe entspricht dem bekannten Multi-Target-Tracking. Eine übliche Herangehensweise ist in Abbildung 2-2 dargestellt und dient in den folgenden Abschnitten als Leitfaden.

In einem Multi-Target-Tracking-System ist der zu schätzende Zustand durch eine Liste von  $N_x$  Objekten gegeben. Da die Objekte Schätzungen des Zustands, der Genauigkeit und der Zuverlässigkeit darstellen, werden sie im Folgenden als Objekthypothesen bezeichnet. Diese werden anhand der  $N_z$  Messungen bzw. Beobachtungen aus einem oder mehreren Sensoren in jedem Messzyklus neu geschätzt. Die üblichen Schritte sind hierbei:

- Synchronisation: Die Beobachtungen und Objekthypothesen müssen auf eine gemeinsame Zeitachse bezogen werden. Die Fusion muss mit unregelmäßigen Sensortakten oder mit der nicht eingehaltenen Reihenfolge der Datenankunft (*out-of-sequence*-Verarbeitung) arbeiten können.
- Prädiktion: Der Zustand der bereits bekannten Objekte muss nach dem ausgewählten Bewegungsmodell von einem vergangenen auf einen aktuellen Zeitpunkt prädiziert werden.
- Alignment/Standardisierung: Die Sensordaten (Beobachtungen) und die Prädiktionen müssen auf ein gemeinsames Koordinatensystem bezogen und in dieselben Einheiten ausgedrückt werden (anhand der Kalibrationsdaten). Hier werden sensorzentrierte, plattformzentrierte oder Weltkoordinaten verwendet.
- Assoziation: Hiermit wird die Zuordnung zwischen Beobachtungen und Prädiktionen bezeichnet. Dafür wird eine Metrik zwischen einer Beobachtung und einer Prädiktion definiert (wie Abstand oder Ähnlichkeit) und die Kombination ausgewählt, die sie minimiert. Da-



**Abbildung 2-2:** Prozessmodell eines Multi-Target-Tracking-Systems.

zu kann die Distanz pro einzelne Beobachtung (Iterativer *Nearest-Neighbor*-Algorithmus, *Strongest-Neighbor*-Algorithmus) oder gleichzeitig für alle Beobachtungen minimiert werden (Globale *Nearest-Neighbor*-Algorithmen, wie *Brute-Force*, Munkres, Auktionsalgorithmen und andere).

Bei Bayes'schen Herangehensweisen wird die Wahrscheinlichkeit jeder möglichen Zuweisung zwischen Beobachtungen und Prädiktionen minimiert, wie das PDAF<sup>3</sup>, JPDAF<sup>4</sup> (vgl. [11]) oder Dempster-Shafer-Verfahren. Nach Klaus [84] können die Assoziationsverfahren einzelne oder mehrfache Zuweisungen pro Beobachtung bzw. Objekthypothese (Modellobjekt) erlauben. Weitere Möglichkeiten erlauben es, die Zuweisung über mehrere Messzyklen durchzuführen.

Alternativ zu diesen Assoziationsverfahren kann ein georeferenziertes Gitter (*occupancy grid*) verwendet werden, in dem die Beobachtungen gewichtet mit deren Zuverlässigkeit und Genauigkeit direkt dargestellt werden.

- **Gating:** Zur Unterstützung der Assoziation kann *a priori* Wissen über die Situation verwendet werden. Dabei werden zum Beispiel Einschränkungen in den Bewegungen der Targets definiert (maximale/minimale Geschwindigkeit, Richtungsänderung etc., vgl. Klaus [84]). In der Regel wird die Größe des Gates anhand der Objektunsicherheit gewichtet (vgl. Bar-Shalom und Li [11]). Auch der Vergleich der Ausdehnung der Prädiktionen zu den Beobachtungen kann dazu dienen, geteilte oder verschmolzene Beobachtungen zu vermeiden (*Splitting/Merging*, vgl. Cramer [35]).
- **Zustandsschätzung:** Anhand der eingehenden, assoziierten Beobachtungen wird der Zustandsvektor jeder Objekthypothese durch Anwendung eines Schätzverfahrens aktualisiert. Die Bayes'sche Schätzmethode für lineare, dynamische Systeme mit weißem Gauß'schem Rauschen ist der Kalman-Filter (KF). Varianten davon sind der Informationsfilter (IF), der Alpha-Beta-Filter, für kontinuierliche Zeit der Kalman-Bucy-Filter, für nichtlineare Systeme der erweiterte Kalman-Filter (EKF), der erweiterte Informationsfilter (EIF) oder der *unscented* Kalman-Filter (UKF). Andere adaptive Strategien sind das adaptive Prozessrauschen (*fudge factor*), Input-Estimation, Variable-State-Dimension oder Multi-Modell-

<sup>3</sup>PDAF – Probabilistic data association filter.

<sup>4</sup>JPDAF – Joint probabilistic data association filter.

Filter (*interacting multi-model*, IMM), vgl. [10].

Im Falle von nichtlinearen Systemen oder nicht-Gauß'schem Rauschen kann anstelle dieser Filter ein Partikelfilter (auch Sequenzielle Monte-Carlo-Methode, SMC) Anwendung finden. Diese Methode rechnet die Wahrscheinlichkeitsdichte des Zustandsraums (*Likelihood*) anhand von einer Wolke von Partikeln (Punkte im Zustandsraum) aus, deren Gewichtungen sequenziell mit den Beobachtungen angepasst werden. Nachteilig in dieser Methode ist der höhere Rechenaufwand der Simulationen. Weitere Methoden sind beispielsweise in den Bereichen der neuronalen Netze, der Fuzzy-Logik oder der Evidenztheorie von Dempster-Shafer zu finden.

Einige Literaturquellen zu Schätzmethoden sind Bar-Shalom und Li [10], Brown [23], Maybeck [106], Niemeier [112]. Details zum Kalman-Filter und dem Informationsfilter werden in Kapitel 7 und in Anhang C gegeben.

- Verwaltungsaufgaben: Hiermit sind Mechanismen gemeint, die die Initialisierung neuer Objekte (z.B. aus nichtassoziierten Beobachtungen), das Löschen veralteter Objekte (z.B. aus Mangel an Beobachtungen oder wegen unnatürlicher Bewegung) sowie das *Splitting* oder *Merging* von Objekten behandeln.

## 2.3 Objekterkennung aus einzelnen Sensoren

Im Fokus dieser Recherche liegen fahrzeugseitig verbaute Sensoren, womit die infrastrukturbaasierten oder kommunikationsbasierten Systeme eine Nebenrolle einnehmen. Außerdem werden hauptsächlich Fahrzeug- und allgemeine Objekterkennungssysteme erläutert, womit auf spezialisierte Erkennungssysteme wie beispielsweise eine Fußgänger- oder Verkehrszeichenerkennung nicht weiter eingegangen wird.

Die Tabelle 2-1 auf Seite 16 stellt einen Gesamtvergleich der Eigenschaften der in diesem Abschnitt betrachteten Sensoren mit Betonung ihrer Stärken und Schwächen vor.

### 2.3.1 Radar

Radarsensoren sind aktive Sensoren, die sich durch eine hohe Entfernungsgenauigkeit aber eine beschränkte Winkelschätzung bzw. Trennbarkeit auszeichnen. Radarsysteme liefern ausschließlich Reflexionspunkte von Objekten mit einer speziellen Signatur, die für den Spektralbereich charakteristisch ist (bspw. metallische Objekte). Bei Radardetektionen liegt keine Korrelation zur Größe, Form und Art der Objekte vor. Sie sind robust gegen Licht- und Wetterverhältnisse, mit Ausnahme von starkem Niederschlag. Durch Nutzung des Doppler-Effektes können Radarsensoren sehr genaue Aussagen über die Geschwindigkeit der Objekte treffen. Die Erkennbarkeit eines Gegenstandes hängt von dessen Reflexionseigenschaften ab. Öffnungswinkel, Reichweite und Anzahl von erkennbaren Objekten hängen von der Art des Radars ab. Als übliche Standardbauformen arbeiten *Short Range Radars* (SRR) im 24 GHz-Bereich und *Long Range Radars* (LRR) im 77 GHz-Bereich (siehe Tabelle 2-1 auf Seite 16).

Da Radarsensoren relativ kostengünstig und klein sind, werden sie häufig in Serienprodukten verwendet, wie beispielsweise bei einem ACC. Jordan et al. [77] führt eine Hindernisklassifikation mit einem LRR basierend auf der Evidenztheorie von Dempster und Shafer durch. Schiementz

[128] fusioniert die Daten mehrerer Radare um ein Fahrzeug. Dabei beschäftigt er sich mit unterschiedlichen Assoziations- und Schätzverfahren, auch unter Berücksichtigung der Objektausdehnung.

### 2.3.2 Lidar und mehrstrahliges Lidar

Ein Lidar ermöglicht die aktive Messung des Abstandes in einer Richtung. Durch parallele Verwendung von mehreren Strahlen (Multi-Beam Lidar) kann dieser erweitert werden. Dabei bleibt die Winkelauflösung gering. Lidars sind robust gegen Licht- und Wetterverhältnisse, mit Ausnahme von dichtem Nebel. Die weiteren Eigenschaften sind der Tabelle 2-1 zu entnehmen.

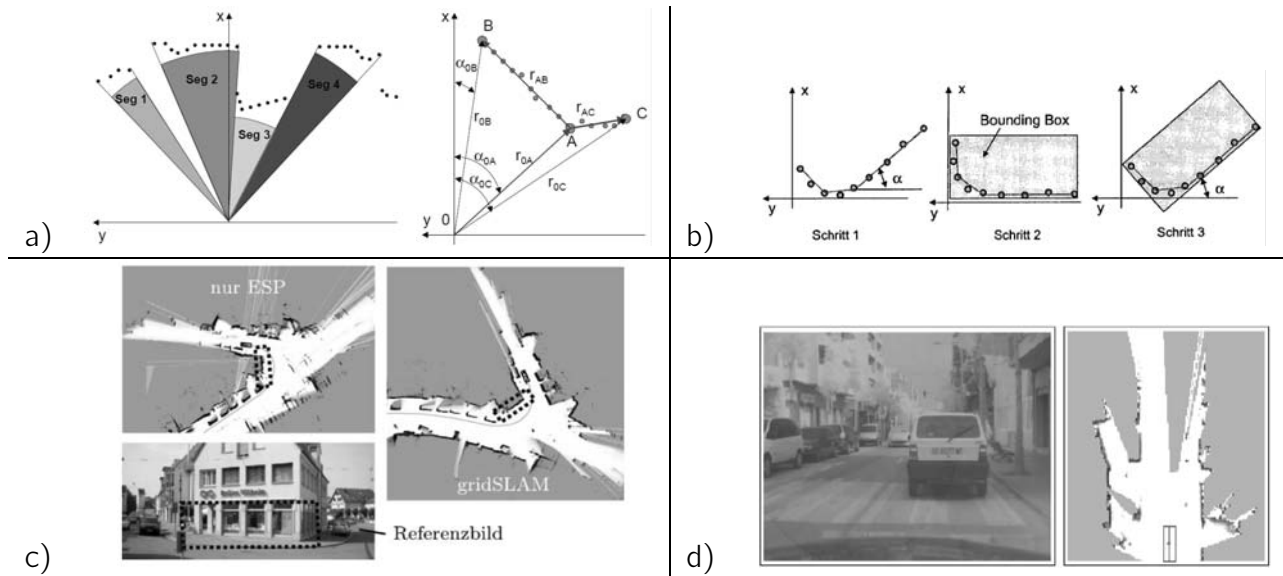
Beispiele für die Verwendung von Lidars, fusioniert mit Kameradaten zur Objekterkennung, finden sich bei Mählich, Dietmayer et al. in [101] und [99].

### 2.3.3 Laserscanner

Laserscanner sind ebenso aktive Sensoren. Durch zusätzliche optische Anordnungen (Drehspiegel) funktionieren sie als Profils Scanner. Daher zeichnen sie sich durch einen großen Erfassungsbereich aus. Die Messung der Entfernung ist dabei sehr genau. Wie Lidars sind sie nur durch dichten Nebel eingeschränkt. Ein Nachteil ist das Auftreten von *false positives* bzw. Phantomobjekten, die durch Bodenreflexionen erzeugt werden können. Die Verarbeitung der Daten eines Laserscanners schließt häufig das Clustern der Punktwolken (Segmentierung), das Ziehen von Konturen um die Objekte und deren Klassifikation mit ein. Abbildung 2-3 stellt Beispiele von einigen der folgenden Systeme dar.

Die Gruppierung bzw. Segmentierung der vom Laserscanner erfassten Messpunkte wird üblicherweise basierend auf dem Abstand zwischen den Punkten durchgeführt. Der Schwellwert des erlaubten Abstands zwischen Punkten eines Objekts wird nach Entfernung und Geschwindigkeit der Objekte angepasst – Vergleiche Dietmayer et al. [42] (Abbildung 2-3.a), Fürstenberg und Dietmayer [59], Wender et al. [161], Streller und Dietmayer [143] und Mendes et al. [107]. Selbst nach der Segmentierung weisen Laserscannerdaten in Fahrzeuganwendungen oft und insbesondere in den Straßenseitenbereichen ein starkes Rauschen auf. Dieser Effekt ist durch die unklaren Konturen (Vegetation am Straßenrand), die Eigenbewegung sowie die bei manchen Sensoren geringere Auflösung in diesen Bereichen bedingt. Außerdem treten häufig bei größeren bzw. entfernten Objekten *split*-Effekte und bei nah beieinander befindlichen Objekten *merge*-Effekte auf.

In den weiteren Schritten werden die Segmente an die ausgewählten Modelle angepasst. Dabei ist der sich verändernde beobachtbare Teil der Objekte im Laufe der Zeit oder durch Verdeckungen und das Fehlen eines festen Referenzpunktes zu berücksichtigen. Dafür verwenden einige Autoren ein Punktmodell, wie Skutek [135] und Labayrade et al. [89, 88]. Andere verwenden ein Zwei-Linien-Modell, wie Fürstenberg und Dietmayer [59] und Dietmayer et al. [42], auch als Drei-Punkte-Verfahren bezeichnet (Becker [13]). Letztere trennen die Objektkontur in drei Punkte, und zwar die Endpunkte und den Punkt, der am nächsten zum Ego-Fahrzeug liegt (erkennbar in Abbildung 2-3.a). Andere Autoren verwenden vordefinierte Formen wie I-, L- und O-Formen, wie Wender et al. [161], ein achsenparalleles Quader, wie Floudas et al. [51], ein ausgerichtetes



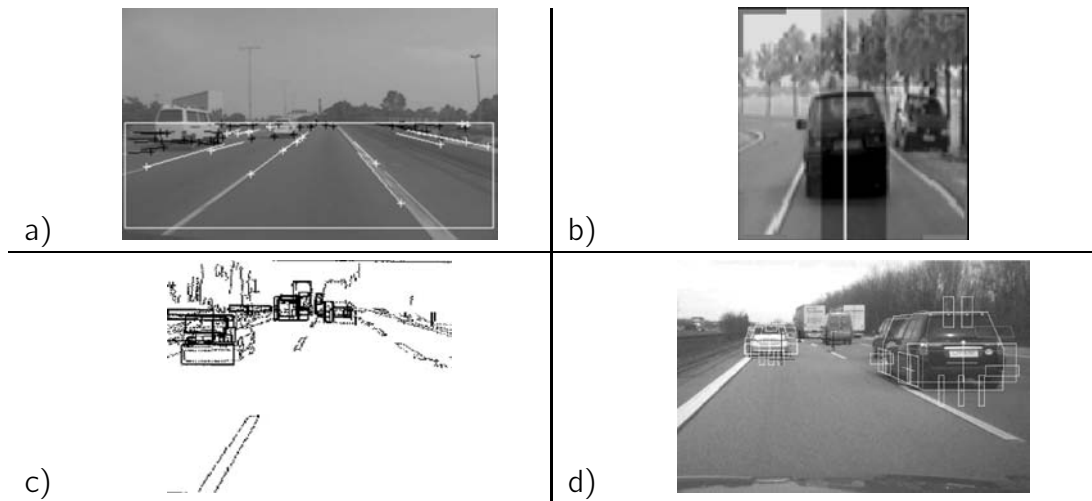
**Abbildung 2-3:** Beispiele von Laserscannerverarbeitung zur Objekterkennung. a) Segmentierung und Drei-Punkte-Modell nach Dietmayer et al. [42], b) Bounding-Box nach Streller et al. [143], c) *Occupancy grid* von Weiß et al. [157, 156, 158] und d) von Vu et al. [154].

Quader mit vordefinierten Größen, wie Streller und Dietmayer [143] (Abbildung 2-3.b), eins mit angepassten Größen, wie der *rotating-calipers*-Algorithmus von Lindner et al. [92], oder sensornahe Modelle für jeden Objekttyp, wie Wender et al. [162] und Cramer [35]. In der Regel haben diese Methoden die Einschränkungen, dass sie für runde oder leicht gerundete Objekte und für konkave Objekte nicht geeignet sind. Außerdem gehen die Methoden davon aus, dass nur die zum Sensor hingewendete Seite erfasst wird. Der Laserscanner kann jedoch auch Punkte unter den Objekten erfassen.

Laserscannerdaten können auch mittels eines georeferenzierten Gitters (*occupancy grid*) verarbeitet werden, wie von Stiller et al. [142, 141], Weiß, Dietmayer et al. [157, 156, 158] (Abbildung 2-3.c), Nieto et al. [113], Vu, Aycard et al. [154] (Abbildung 2-3.d), Wender, Dietmayer et al. [162], Lindner et al. [92] und Skutek [135] vorgestellt.

### 2.3.4 Mono-Bildverarbeitung

Die Extraktion von Informationen aus einem Kamerabild ist in der Regel komplex (Quellen zur Bildverarbeitung: Russ [125], Jähne [75], Steinbrecher [138], Bradski und Kaehler [18]). Grundsätzliche Probleme der Bildverarbeitung für die Objekterkennung im Außenbereich sind die Abhängigkeit von Lichtverhältnissen und Witterung, sowie das Auftreten von Verdeckungen, Schatten und Artefakten. Die Genauigkeit der visuellen Objekterkennung hängt stark von der Entfernung zum Objekt ab. Zur Einschätzung der Entfernung treffen diese Methoden Annahmen, wie zum Beispiel eine flache Straßenoberfläche oder eine feste Objektgröße. In Tabelle 2-1 werden Mono-Kamerasensoren im Allgemeinen mit den restlichen Sensoren verglichen. Die grundsätzlichen Herangehensweisen zur Objekterkennung in der Bildverarbeitung können in folgende allgemeine Gruppen eingeteilt werden, die zum Teil exemplarisch in Abbildung 2-4 dargestellt sind: modellbasierte Objekterkennung, bewegungsbasierte Objekterkennung, Erkennung basierend auf dem Objektschatten und merkmalsbasierte Objekterkennung.



**Abbildung 2-4:** Einige Ansätze zur Objekterkennung aus Mono-Kamerabildern. a) Bewegung [166], b) Symmetrie [14], c) Konturlinien [163], d) merkmalsbasierte Ansätze [111, 110].

In der modellbasierten Objekterkennung wird das Bild anhand der Kenntnis über das zu suchende Objekt in Bildkoordinaten abgesucht. Die Objekterkennung kann beispielsweise auf dem Anlernen von Zielobjektypen beruhen, wie bei Viola und Jones [152] anhand von Haar-Wavelet-Merkmalen beschrieben. Der Nachteil der modellbasierten Systeme ist die hohe Abhängigkeit von den Modellen, womit ein Objekt unbekannten Typs (bspw. ein Hindernis mit bisher unberücksichtigten Merkmalen) nicht detektierbar ist.

Die bewegungsbasierten Methoden zur Objekterkennung nutzen anstatt von Modellen die Bewegung im Bild (optischer Fluss). Aus dem optischen Fluss lässt sich die eigene Bewegung der Kamera abschätzen. Dabei können aus Bildregionen mit abweichenden Bewegungsmustern Objekte identifiziert werden, wie Pauwels et al. [114], Klappstein, Franke et al. [83] sowie Giachetti et al. [64] zeigen. Ist die Eigenbewegung anderweitig bekannt, so können die Bilder paarweise als *Motion-Stereo* betrachtet werden<sup>5</sup>, wie bei Kato et al. [81] sowie Wohlfeil und Reulke [166] (Abbildung 2-4.a) gezeigt. Die Nachteile der bewegungsbasierten Ansätze liegen in der Unfähigkeit, statische oder gleichschnell fahrende Objekte zu erkennen. Auch im Fluchtpunktbereich des Bildes sind Bewegungen schwer zu erkennen.

Viele Ansätze zur Objekterkennung treffen die Annahme, dass alle Objekte einen deutlichen Schattenbereich bzw. Schattenzug aufweisen. Beispiele hierzu sind Thomanek [145], Wohlfeil und Reulke [166] und Kämpchen, Dietmayer et al. [80]. Die Einschränkung dieser Ansätze ist, dass der Schatten deutlich sichtbar sein muss, was in vielen Fällen gar nicht oder nur teilweise zutrifft.

In der heterogenen Gruppe der merkmalsbasierten Methoden wird die Objekterkennung anhand bestimmter Merkmale durchgeführt, wie vertikale Kanten, Räder oder Symmetrie. Zusätzlich zu dem Objektschatten basiert die Arbeit von Thomanek [145] auf Schatten und vertikalen Kanten. Die von Wohlfeil und Reulke [166] auf optischem Fluss, und die von Kämpchen, Dietmayer et al. [80] auf vertikale Kanten, Symmetrie und optischem Fluss. Bensrhair, Bertozzi, Broggi et al. [14] (Abbildung 2-4.b) suchen nach Objekten in Bildern anhand von Symmetrie, Kontur und Seitenverhältnis. Kolski et al. [86] stellen ein Multi-Agenten-System basierend auf Symmetrie,

<sup>5</sup>Als kooperative Fusion von Ortung und Bild

**Tabelle 2-1:** Einschätzung von Sensoren bezüglich unterschiedlicher Komponenten der Objekterkennung.

Sensor	Datenrate	Reichweite	Öffnungswinkel	Entfernungs- genauigkeit	Winkelauflösung	Geschwindigkeits- schätzung	Zuverlässigkeit	Objekt-splitting und -merging	Klassifikation	Wetter- empfindlichkeit
Radar (SRR)	++ bis 15 Hz	0 bis 50 m	++ bis $\pm 60^\circ$	++	0	++	++	--	--	-
Radar (LRR)	++ bis 40 Hz	++ bis 150 m	- bis $\pm 10^\circ$	++	-	++	++	--	--	-
Mehrstrahl- iges Lidar	++ bis 20 Hz	++ bis 200 m	0 bis $16^\circ$	++	-	0	+	--	--	-
Laser- scanner	++ 10- 40 Hz	++ bis 200 m	++ bis $360^\circ$	++	+	0	++	-	0	-
Mono- kamera	++ 24 Hz	+ bis 80 m	+ $\sim 30^\circ$	-	++	0	-	0	++	--
Stereo- kamera	++ 24 Hz	+ bis 80 m	+ $\sim 30^\circ$	+	++	0	+	0	++	--
3D-Kamera	++ bis 80 Hz	0 bis 50 m	+ $\sim 50^\circ$	+	+	0	+	--	0	--

Schatten, Rädern und Windschutzscheiben vor. Die Arbeiten von Achler und Trivedi [6, 5] sowie Wender und Dietmayer [159, 160] basieren auf der Suche nach Rädern. Wetzel [163] verwendet Merkmale wie Kanten, Regionen und Geraden zur Objekterkennung (Abbildung 2-4.c). Schließlich verwenden Schwarzingler [130], von Holt [73], Armingol, Collado et al. [8, 34] und Neumaier, Färber et al. [111, 110] (Abbildung 2-4.d) Modellobjekte basierend auf Kanten und Features. Alle diese Methoden beruhen auf der Annahme, dass die zu erkennenden Objekte diese Merkmale besitzen bzw. dass diese im Bild deutlich sichtbar sind. Falls diese Annahme nicht zutrifft, sind die Objekte nicht detektierbar.

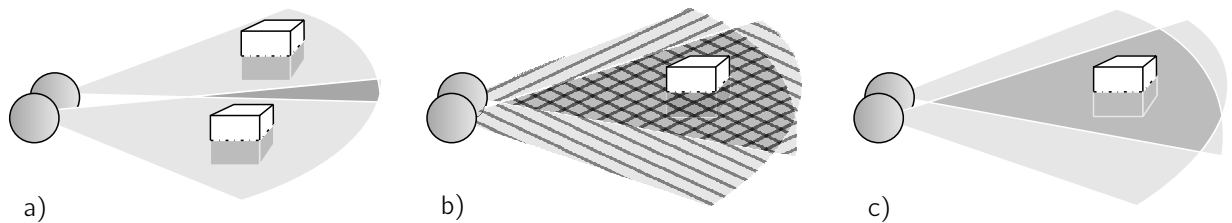
### 2.3.5 Stereobasierte Objekterkennung

Neben den auf einer Mono-Kamera basierten Bildverarbeitungssystemen stellen die Stereo-Systeme eine weitere wichtige Gruppe dar. Der Vorteil der Stereovision ist, dass zur Bildinformation die Tiefeninformation gewonnen wird. Als Nachteile sind die vergleichsweise aufwendige Verarbeitung sowie die Empfindlichkeit in Bezug auf Lichtverhältnisse, Blendung und allgemeine Witterung zu nennen. Einige Arbeiten in diesem Bereich basieren auf einem *sparse-stereo*-Ansatz<sup>6</sup>, wie von Franke et al. [55, 53] und Stiller et al. [140]. Andere verwenden einen *dense-stereo*-Ansatz<sup>7</sup>, wie von Benschir, Bertozzi, Broggi et al. [14], Saneyoshi et al. [126] und Franke [54]. Letzterer kombiniert den Stereoansatz mit einem bewegungsbasierten Verfahren.

<sup>6</sup>In *sparse-stereo*-Ansätzen wird die Tiefeninformation für eine Untermenge von ausgewählten Bildpunkten (*feature*) berechnet.

<sup>7</sup>Stereo-Information wird in *dense-stereo*-Ansätzen für das gesamte Bild berechnet.





**Abbildung 2-5:** Beispiele für Fusionsprinzipien: a) Komplementär zur Erweiterung des Erfassungsbereichs; b) Kooperativ zur Erweiterung der erfassten Informationen; c) Konkurrierend zur Erhöhung der Genauigkeit.

Stereo-Vision kann auch mit einer Orthorektifizierung kombiniert werden, wie von Bertozzi und Broggi in [16] vorgestellt.

### 2.3.6 3D-Kamera

Basierend auf aktiver Lichtmodulation zeichnet eine 3D-Kamera für jedes Pixel neben einem Grauwert auch eine Entfernungsangabe auf. Daher stehen für die Bildsegmentierung analog zu der Stereo-Kamera auch die Tiefeninformationen zur Verfügung. Die Nachteile dieser Kameras sind die aktuell noch beschränkte Bildauflösung sowie der störende Effekt von Sonnenlicht im Außenbereich. Verfahren zur Objekterkennung anhand einer solchen Kamera sind von Fardi et al. [50] und Ringbeck et al. [122] vorgestellt worden.

## 2.4 Fusionsansätze zur Objekterkennung

Wie aus der Aufstellung zu ersehen, ist trotz der Vielfalt an Sensoren und Methoden bei der Objekterkennung immer mit Einschränkungen und Nachteilen zu rechnen. In diesem Abschnitt werden Systeme vorgestellt, die die Aufgabe der Objekterkennung mittels Fusion lösen. Die Vorteile der Fusion wurden in Abschnitt 1.2 bereits erläutert.

Ein grundlegendes Kriterium zur Klassifikation von Fusionssystemen ist das verfolgte Fusionsprinzip<sup>8</sup> (siehe Abbildung 2-5):

- Komplementäre Fusion (*complementary fusion*): Die unterschiedlichen Erfassungsbereiche und/oder die unterschiedliche gewonnenen Informationen der Sensoren werden zusammengeführt. Die Aufgabe der Fusion ist, mit dem erweiterten Bereich umzugehen. Beispiele dazu sind die Abdeckung eines größeren Bereichs durch mehrere Radarsensoren, die Erzeugung von Panoramabildern auf Basis mehrerer Kameras oder die Fusion eines Fern- und eines Nahbereichsensors. Durch dieses Prinzip ergeben sich bis auf den Übergangsbereich oder bei gleichen Merkmalen keine Konflikte. Im Konfliktfall bzw. Übergangsbereich wird eines der weiteren Prinzipien verwendet. Werden keine Querverweise zwischen den Daten genutzt oder gesetzt, so heißt diese Nebeneinanderlegung der Daten unabhängige Fusion (Brooks, [21]).
- Kooperative Fusion (*cooperative fusion*): Die Daten der Sensoren werden so verarbeitet, dass die erfassbaren Informationen zusammengeführt und möglicherweise erweitert werden.

<sup>8</sup>Das Fusionsprinzip wird in der Literatur häufig als Fusionsparadigma genannt.

Damit ergibt sich ein Gewinn in der Genauigkeit, Zuverlässigkeit oder eine neue Information, die keine der einzelnen Sensoren für sich hätte erkennen können, als handele es sich um einen virtuellen Sensor. Diese Fusion ist meistens sehr sensornah und daher weniger verallgemeinerbar. Zum Beispiel wird Stereo-Vision als kooperative Fusion zur Gewinnung von Tiefeninformation über die Szene aus zwei Kamerabildern bezeichnet. Ebenso als kooperativ bezeichnet, ist die Unterstützung oder Validierung der Daten eines Sensors durch einen anderen. Zum Beispiel können Interessenbereiche in einem Kamerabild anhand der Daten eines Laserscanners gesetzt werden.

- **Konkurrierende Fusion (*competitive fusion*):** Die Daten von gleichen (homogen) oder unterschiedlichen (heterogen) Sensoren geben Aufschluss über denselben Informationsinhalt (eventuell nach Vorverarbeitung) im selben Beobachtungsbereich. Durch konkurrierende Fusion erfolgt ein Gewinn in der Genauigkeit und/oder Zuverlässigkeit. Dabei können oft Konflikte in Form von widersprüchlichen Informationen auftreten, die gelöst werden müssen. Solche Widersprüche lassen sich beispielsweise mit Rückgriff auf die Rohdaten der Sensoren lösen. Dieses Fusionsprinzip wird oft auf der mittleren Abstraktionsebene verwendet, da die Daten jedes Sensors vorverarbeitet werden. Sind die Sensoren und deren Erfassungsbereiche identisch, so spricht man von einer redundanten Fusion<sup>9</sup>.

Die Entscheidung, ob die Sensordaten kooperativ oder konkurrierend fusioniert werden sollten, ist nicht immer leicht zu treffen. So können nach Brooks [21] zwei leicht unterschiedliche Beobachtungen eines Objektes kooperativ seine Größe und konkurrierend seine Positionsungenauigkeit ergeben. In einem Multi-Target-Szenario können die beiden Beobachtungen auch als zwei unterschiedliche Objekte betrachtet werden.

Fusionssysteme werden meist nicht nur nach einem dieser Prinzipien entwickelt, sondern besitzen im Allgemeinen mehrere der erläuterten Eigenschaften. Daher spricht man hier von einer hybriden Fusion.

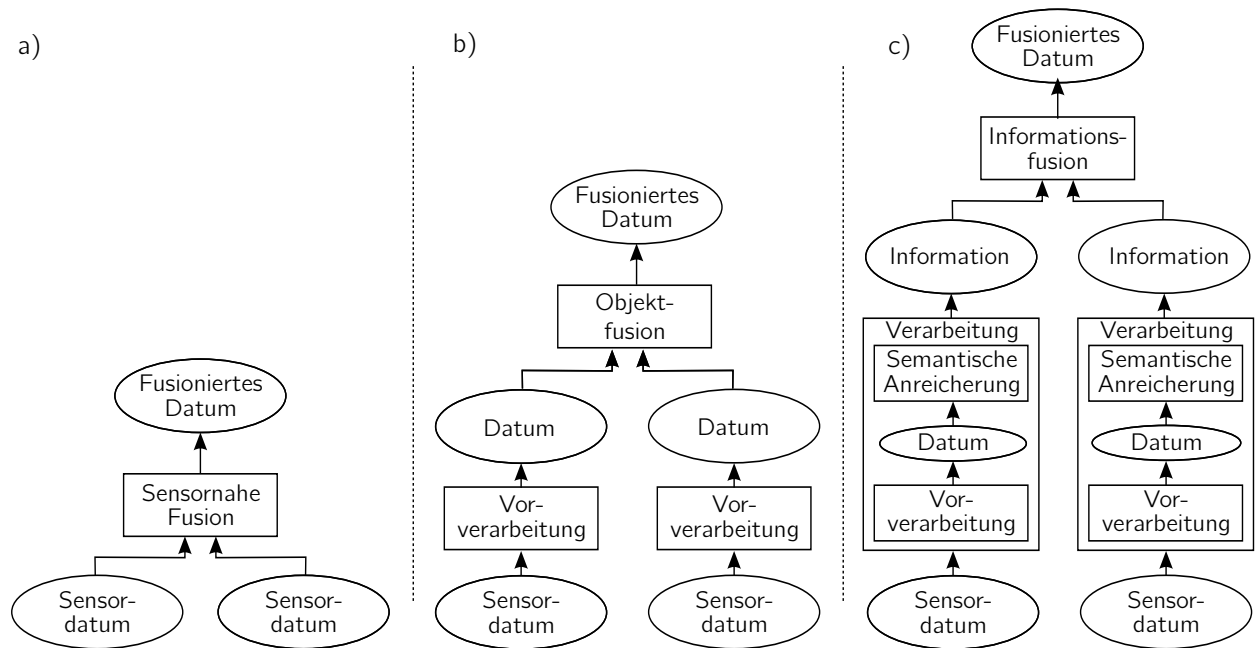
In dieser Literaturrecherche werden die Systeme zur Objekterkennung nach deren Abstraktionsebene klassifiziert (siehe Abbildung 2-6):

- **Sensornahe Fusion (*low level*):** Die Sensorrohdaten werden direkt miteinander kombiniert.
- **Objektfusion (*mid level*):** Die Fusion folgt einer vorhergehenden Verarbeitung, wodurch die Menge an Daten auf Merkmale oder Objekte reduziert ist. Wenn die Objekte im Rahmen des Trackings fusioniert werden, spricht man von Tracking-Fusion.
- **Informationsfusion (*high level*):** Die Daten werden zunächst auf eine semantische Ebene gebracht und erst dann fusioniert. Diese semantische Ebene wird auch Entscheidungs-, Anwendungs- oder Identifikationsebene genannt.

Diese Einteilung hängt eng mit den vorhin erwähnten Fusionsprinzipien zusammen. Auch hier verschwimmen die Grenzen zwischen den Gruppen miteinander. In einer sensornahen Fusion werden häufig kooperative Ansätze verwendet. Der Nachteil dieser Herangehensweise ist, dass sie sehr von den konkreten Sensoren abhängig ist und oft Eingriffe und Verschmelzungen in

---

<sup>9</sup>Unter Redundanz versteht sich auch das Vorhandensein von Notsystemen, die beim Ausfall des Hauptsystems eingeschaltet werden können. Dies liegt nicht im Bereich der Fusion.



**Abbildung 2-6:** Abstraktionsebenen der Fusion: a) eine sensornähe Fusion, b) eine Objektfusion und c) eine Informationsfusion.

der Datenverarbeitung der Einzelsensoren umfasst (z.B. in der Bildverarbeitung). Dafür ist aber eine sensornähe Fusion üblicherweise sehr leistungsfähig und problemoptimiert. Auf Objekt- oder Trackingebene wird häufig ein konkurrierender Ansatz verwendet. Je abstrakter die Fusion ist, desto allgemeiner sind die dabei verwendeten Methoden und Techniken. Somit haben sie eine höhere Zuverlässigkeit, sie sind einfacher zu erweitern und Komponenten können leichter ausgetauscht werden. Da abstrakte Fusionssysteme in der Regel auch komplexer sind, müssen dabei Abstriche im Bereich der Geschwindigkeit (wegen mehr Overhead) oder Leistung (weniger Information als in den Rohdaten) hingenommen werden.

Fusionssysteme werden oft aufeinander aufbauend in so genannten Multi-Level-Sensorfusionssystemen erstellt.

Tabelle 2-2 stellt einen Vergleich zwischen den Objekterkennungssystemen auf den jeweiligen Ebenen dar. Hierbei wird deutlich, dass jede Ebene Vor- und Nachteile aufweist. In folgenden Abschnitten werden Details und Beispiele aus der Literatur vorgestellt. Abbildung 2-7 zeigt Beispiele für einige der erwähnten Systeme. Alle diese Systeme benutzen das in Abschnitt 2.2 vorgestellte Modell des Multi-Target-Tracking.

### 2.4.1 Sensornähe Fusion

Auf Sensordatenebene gibt es Systeme, die die Daten einer Kamera und eines Laserscanners fusionieren, wie zum Beispiel Kämpchen, Fürstenberg und Dietmayer [80, 79] (Abbildung 2-7.a), Wender und Dietmayer [159, 160], Neumaier und Färber [110] (bereits in Abbildung 2-4.a gezeigt). Andere fusionieren die Daten einer Kamera und eines Radars, wie zum Beispiel Bombini et al. [17], Kato et al. [82], Shimomura et al. [132], Beauvais und Lakshmanan [12]. Dabei führt die Kamera eine Verfeinerung [80, 79][81][12], Erkennung von *false positives* [17], Klassifikation [80, 79][132] oder Verfolgung [110] der Objekte durch. In [79] wird ein auf dem

**Tabelle 2-2:** Vergleich der grundsätzlichen Ansätze zur Objekterkennung anhand von Sensorfusion

Fusionsansatz	Performanz	Verfügbarkeit/ Sensorabhängig	Kopplungsgrad	Skalierbarkeit/ Erweiterbarkeit	Verteilbarkeit	Genauigkeit	Zuverlässigkeit
Sensorfusion (unterstützend)	++	--	--	--	--	++	+
Sensorfusion (gitterbasiert)	++	+	++	++	+	-	++
Objektfusion (Merkmale)	+	0	-	0	+	+	++
Objektfusion (Objekte)	+	+	++	+	+	+	++
Tracking-Fusion	--	++	+	++	++	+	-

Kalman-Filter basierter *Interacting Multi-Model-Filter* verwendet (IMM). Der Laserscanner bzw. Radar ist bei diesen Systemen der Hauptsensor. Fällt dieser aus, so gilt dies auch für das gesamte Fusionssystem, womit die Verfügbarkeit gefährdet ist. Die Methoden der Fusion umfassen die Transformation der Daten des Laserscanners bzw. Radars in das Kamerabild und die Anwendung von Methoden der Bildverarbeitung, wie in Abschnitt 2.3.4 vorgestellt wurde.

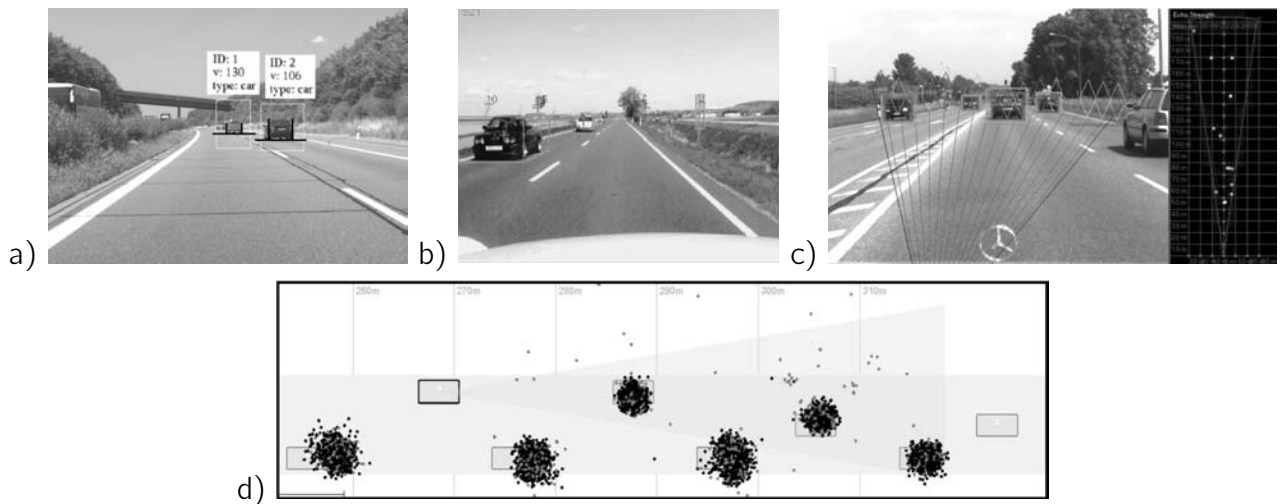
Andere Autoren verwenden ein georeferenziertes Gitter (*occupancy grid*) für die sensornahe, konkurrierende Fusion von Daten. Im Gitter können die Sensordaten überlagert werden, womit das Problem der Assoziation entfällt. Beispiele dazu sind Aycard [9] (Abbildung 2-3), Skutec und Linzmeier [136] und Kolski et al. [86]. Die Performanz und Genauigkeit der gitterbasierten Ansätze hängen von der Gitterauflösung ab.

## 2.4.2 Fusion auf Objektebene

Jeder Sensor führt eine mehr oder weniger aufwändige Vorverarbeitung der Daten durch und die erzeugten Daten werden konkurrierend fusioniert. Darin gibt es die Unterscheidung, ob die Objektdaten auf Merkmalsebene (auch *early fusion* genannt) oder als Objektbeobachtungen (die eigentliche Objektfusion) assoziiert werden.

Bei der Fusion auf Merkmalsebene geschieht die Assoziation zwischen sensornahen Merkmalen, wie zum Beispiel Kanten und Reflexionspunkten. Dafür werden in jedem Zyklus die Modelldaten auf Sensormerkmale prädiert und mit den aus den Sensordaten extrahierten Merkmalen verglichen. Beispiele für diese Gruppe sind Walchshäusl et al. [155], von Holt [73] und Cramer [35]. Für die Assoziation auf Merkmalsebene wird in [73] das Konzept der Hopfield-Netze verwendet.

Für die Objektfusion werden Objektbeobachtungen abgeleitet von den Daten aus jedem Sensor erzeugt. Die Assoziation findet auf Objektebene statt. Stüker [144] fusioniert eine Kamera und ein Laserscanner auf Objektebene und fokussiert seine Arbeit auf die Verarbeitung asynchroner Sensordaten sowie auf die variable Beobachtbarkeit von Zustandsgrößen. Stiller et al. [140]



**Abbildung 2-7:** Beispiele von Fusionssystemen zur Objekterkennung und Verfolgung. a) Sensornah [80, 79], b) Objektfusion [140], c) JIPDA [102, 100], d) Partikelfilter [123].

(Abbildung 2-7.b) fusionieren die Objektdaten aus einem Stereo-System und einem Laserscanner in einem Kalman-Filter. Becker [13] beschäftigt sich mit der Fusion eines Stereo-Systems, vier Laserscanner und eines Radars. Dabei stellt er für die Assoziation ein Munkres-Verfahren und einen Nearest-Neighbor-Algorithmus vor. Zur Fusion stellt er einen Informationsfilter vor, der sich für wechselnde Sensoranzahl eignet. Zudem erkennt er manövrierende Objekte, um die Systemunsicherheit anzupassen. Gruyer et al. [68, 69] und Vercauteren und Wang [151] schlagen dezentralisierte Fusionssysteme vor. Effertz [45] realisiert die Objekterkennung in Form eines hybriden Ansatzes einerseits mit einer objektbasierten Assoziation und einem Kalman-Filter und andererseits mit einem georeferenzierten Gitter und einer Dempster-Shafer-Fusion. Für weitere Beispiele sei auf Fürstenberg und Rössler [60], sowie Polychronopoulos et al. [115] hingewiesen.

Die erwähnten Beispiele basieren auf dem Kalman-Filter oder einer seiner Varianten (siehe Abschnitt 2.2). Alternativ dazu wird in [131] von Schweiger et al. und in [123] von Röckl et al. über einen Partikelfilter fusioniert (Abbildung 2-7.d). Zuletzt ist auch die Arbeit von Mählich, Dietmayer et al. [102, 100] zu nennen (Abbildung 2-7.c), in dem ein JIPDA<sup>10</sup> Verwendung findet. Diese Methode vermeidet die frühe Entscheidung im Erkennungsprozess und betrachtet die Existenzwahrscheinlichkeit eines Objektes unabhängig von dessen Genauigkeit.

### 2.4.3 Fusion auf Tracking-Ebene

Werden Objekttracks (Trajektorien) statt Objektbeobachtungen verwendet, wird von Tracking-Fusion gesprochen. Eine solche Fusion ist weniger häufig, da hier für jeden Sensor ein vollwertiges Tracking notwendig ist. In dieser Ebene können Fehler schwer herausgefiltert oder nachvollzogen werden, was eine eingeschränkte Zuverlässigkeit nach sich zieht. Manche Systeme können aufgrund von Kommunikations- oder Organisationsstrukturen nicht auf einer tieferen Ebene fusionieren. Dafür sind diese Systeme sehr modular. Diese Ansätze sind in den Arbeiten von Chong et al. [33], Floudas, Polychronopoulos et al. [51] zu finden. Simon und Becker [133] und Becker [13] schlagen auch einen Ansatz zur Fusion auf dieser Ebene vor.

<sup>10</sup> JIPDA – Joint Integrated Probabilistic Data Association.

#### 2.4.4 Informationsfusion

In der nächsthöheren Abstraktionsebene werden Objekte zu Hindernissen, Fußgängern oder Fahrzeugen, deren Manövern, Absichten, Bahnplanung, Verkehrsregeln und im Allgemeinen zu möglichen Gefährdungen in der Fahrsituation assoziiert. In dieser Arbeit wird der Fokus auf potentielle Gefahren gelegt. Die Erkennung von Gefahren wird im nächsten Abschnitt näher erläutert.

### 2.5 Erkennung von Gefahrensituationen

Potentielle Gefahren sind eines der wichtigsten Merkmale, die Einfluss auf das Fahrverhalten haben. Die frühzeitige Erkennung von potentiellen Gefahrensituationen spielt demnach in Assistenz- und Automationssystemen eine wesentliche Rolle. In diesem Abschnitt werden die wichtigsten Begriffe im Zusammenhang mit der Gefahrenerkennung erläutert sowie Ansätze aus der Literatur vorgestellt.

Nach DIN VDE 31000 [149] ist eine Gefahr *eine Sachlage, bei der das Risiko größer als das Grenzkisiko ist* definiert, wobei Risiko als *Zusammenhang zwischen einem Schaden und einer Eintrittswahrscheinlichkeit* und Grenzkisiko als *allgemein akzeptierte Gefahr* definiert sind<sup>11</sup>.

Der Schaden ist im Straßenverkehr als Unfallfolge zu verstehen. Ein Unfall ist nach Strafrecht [24] wie folgt definiert: *Ein Unfall ist jedes plötzliche, mit dem Straßenverkehr und seinen Gefahren ursächlich zusammenhängende Ereignis, das einen nicht völlig belanglosen Personen- oder Sachschaden zur Folge hat*. Diese Definition beinhaltet auch die Ursache-Wirkungs-Beziehung zwischen Gefahrensituation und Schaden.

Zur Analyse von Gefahren und Unfällen kann die Gefahrenträgertheorie [134] (aus dem Gebiet der Arbeitssicherheit) herangezogen werden. Darin ist ein Unfall als *plötzliches, ungewolltes eine Körperverletzung bewirkendes Ereignis* definiert. Dieses kann sowohl vom Subjekt (Mensch) als auch vom Objekt (Gegenstand) ausgehen, da beide als Gefahrenträger gesehen werden (siehe Abbildung 2-8.a). Zur systematischen Analyse von Unfällen gehören die folgenden drei Ebenen, in denen die Bereiche des Menschen, der Organisation und der Technik in Zusammenspiel kommen:

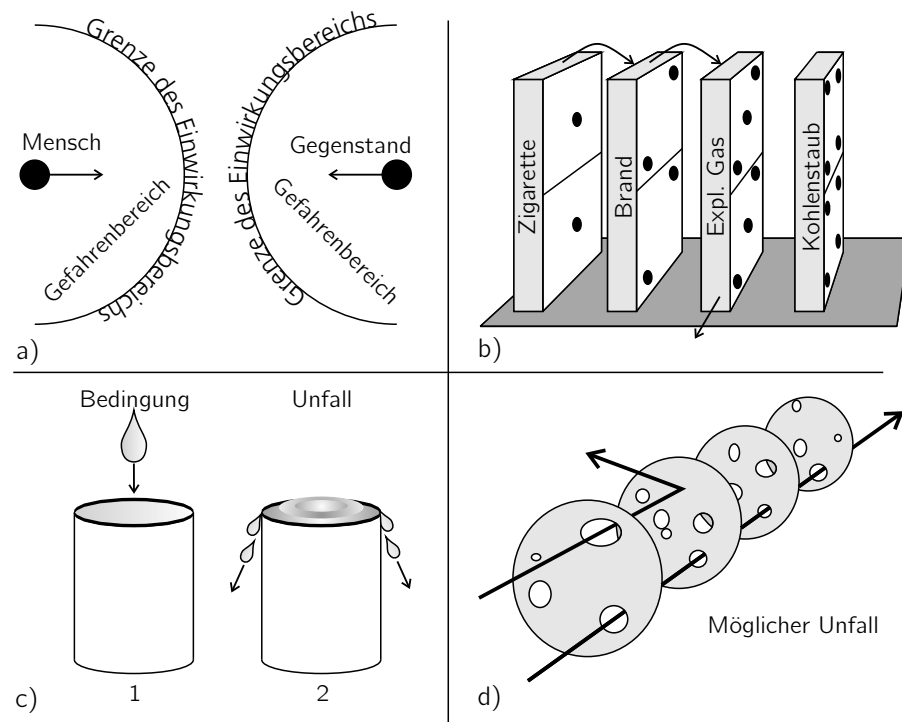
- Unfallbedingungen, die vom Sollwert abweichen und dadurch Unfälle verursachen.
- Unfallhergang.
- Folgen, einschließlich Personen- und Sachschaden.

Aus der Gefahrenträgertheorie lässt sich also schließen, dass ein Unfall durch Abweichung des erwarteten Verhaltens von Menschen, Organisation oder Technik verursacht wird. Zusätzlich betont diese Theorie, dass diese Abweichungen durch viele anderen Faktoren beeinflusst sind, wie der Zustand der Menschen, der Organisation und der Technik.

Auf den Automotive-Bereich übertragen bedeutet das, dass die Ursachen eines Unfalls in einem fehlerhaften Fahrverhalten (nicht Einhaltung der Straßenverkehrsordnung), einer Kontrollverlust, in einem technischen Fehler im Fahrzeug, oder in einer fehlerhaften Infrastruktur (z.B. falsch beschildert oder gebaut) zu finden sind. Der Fahrerzustand, Straßenzustand und Fahrzeugzustand

---

<sup>11</sup>Vergleiche mit Risikobewertungstheorie, [104].



**Abbildung 2-8:** Konzept von Gefahr nach unterschiedlichen Theorien: a) Gefahrsträgertheorie, b) Domino-Effekt, c) überlaufendes Fass und d) Schweizer-Käse-Modell.

spielen daher eine wichtige Rolle. Für die Gefahrenerkennung müssen die Sollzustände definiert sein.

Wie im nächsten Kapitel gezeigt wird, beschäftigen sich viele Assistenz- und Automationssysteme in Fahrzeuganwendungen mit der Erkennung solcher Abweichungen, um somit Fahrurfälle zu vermeiden.

Es gibt andere Modelle, um den Zusammenhang zwischen einer Gefahr und einem Unfall darzustellen. Zum Beispiel können Gefahren in Form einer Kettenreaktion Unfälle produzieren (Metapher: Domino-Effekt [134], Abbildung 2-8.b). Eine andere Metapher sieht die Situation als Fass, das Tropfen für Tropfen mit Gefahren aufgefüllt wird, bis der letzte Tropfen es zum Überlaufen bringt und den Unfall auslöst ([134], Abbildung 2-8.c). Ein weiteres Gefahrenmodell stellt das als Schweizer-Käse bezeichnete Modell dar (Reason [119], Abbildung 2-8.d). Dieses besteht aus verschiedenen Ebenen (Scheiben) mit Löchern, welche jeweils die Barrieren bzw. die Schwächen in der Situation darstellen. Nur bei bestimmten Anordnungen kann die Gefahr bis zum Unfall führen. All diese Gefahrenmodelle sind mit der klassischen Ursache-Wirkungs-Beziehung nur schwer darstellbar, denn viele dieser Ursachen führen nur in bestimmten Kombinationen oder Mengen zum Unfall.

Da diese Kombinationen verschiedener Umstände letztlich selten vorkommen, besteht eine andere Strategie zur Gefahrenerkennung in der Feststellung von atypischen Ereignissen, die in bestimmten Fällen Gefahren darstellen können. Diese Ansicht geht von der an sich geringen Eintrittswahrscheinlichkeit des Schadens aus. Diese Idee findet sich zum Teil auch in den davor erläuterten Theorien wieder. In der Gefahrsträgertheorie ist sie als Abweichung des erwarteten Verhaltens gegeben. Im Schweizer-Käse-Modell ist sie als seltene Anordnungen von Löchern (atypisch) gegeben. Im nächsten Kapitel werden Ansätze vorgestellt, die diesem Weg folgen.

In folgenden Abschnitten werden Systeme vorgestellt, die die Gefahrenerkennung auf Basis einer der gerade erwähnten Strategien durchführen, nämlich als Teil einer Ursache-Wirkungs-Beziehung und als atypische Situation.

### 2.5.1 Gefahren als Ursache-Wirkungs-Beziehung

Zu dieser Gruppe gehören die meisten aktuellen Assistenz- und Automationssysteme, die sich jeweils in der Regel auf eine Gefährdungsart fokussieren. Einige davon wurden in Kapitel 1 eingeführt. Beispiele sind ein LDW, ein NBA, eine Kollisionswarnung (wie von Dagan et al. vorgeschlagen [37]), Kollisionsvermeidung (Jansson [76], Franke et al. [56], Brandt et al. [19]), Kollisionsfolgenverminderung (Labayrade [88]), Kreuzungsassistent ([1], Chaufton [2]), oder ein Toter-Winkel-Warner (Steffel et al. [137]).

Um einen größeren Umfang der Situation zu berücksichtigen, haben andere Ansätze Indikatoren für den Grad an Gefahr einer Situation definiert. Beispiele dazu sind das *Time-to-Collision* (TTC) und *Time-Headway* (THW). Polychronopoulos, Tango et al. [116] schlagen auch die *Predicted Minimum Distance* (PMD) und die *Time to Predicted Minimum Distance* (TPMD) vor.

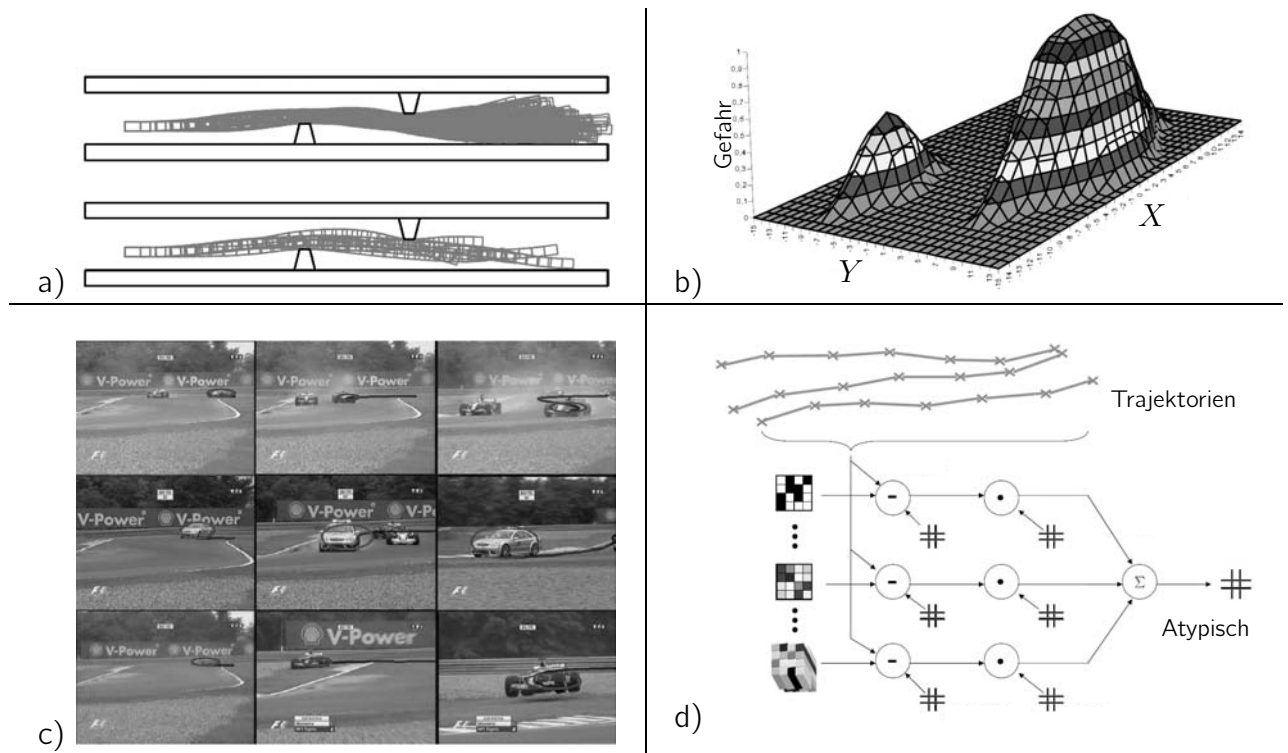
In diesen Systemen sind in der Regel die Schwellwerte festgelegt, bei denen sie warnen oder eingreifen. Dieses ist deswegen ungünstig, weil sich solche Schwellwerte an die Gesamtsituation anpassen sollten. Diese Problematik wird von Yang et al. [169] diskutiert. Als Lösung wird die online Berechnung des optimalen Schwellwertes anhand der Monte-Carlo-Methode vorgeschlagen.

In den Arbeiten von Eidehall [46, 48, 47] (Abbildung 2-9.a) werden mehr semantische Informationen berücksichtigt, wie beispielsweise Objekttyp, die Zuordnung zu einem Fahrstreifen, das durchgeführte Fahrmanöver oder gar die Fahrerintentionen. Die Gefahrenerkennung besteht darin, diese Informationen in eine Monte-Carlo-Simulation einzubeziehen, um die Eintrittswahrscheinlichkeit einer Kollision (Gefahr) zu berechnen. Die Methode benötigt allerdings so viel abstrakte Information der Umgebung, dass sie bisher nur mit simulierten Daten implementiert und demonstriert werden konnte. Die Arbeiten von Broadhurst, Kanade et al. [20] verfolgen ebenfalls diese Vorgehensweise.

Einen weiteren Ansatz zur Gefahrenerkennung bilden die Gefahrenkarten oder Potentialfelder. Diese Karten (meist lokal und egofahrzeugzentriert) beinhalten Informationen über den Grad an Gefahr jeder Position (ähnlich dem Potential in einem elektrischen Feld) aber nicht über deren Wahrscheinlichkeit. Gefährdungen sind darin als Gebirge dargestellt. Hindernisse weisen beispielsweise einen Gefahrenberg um sich herum auf, da in seiner Nähe zu sein, gefährlich wäre (wieder im Sinne der Ursache-Wirkungs-Beziehung). Ähnliche Konzepte wurden bereits in den 1930er Jahren beschrieben, wie von Gibson et al. [65].

Gefahrenkarten findet man oft zur Bahnplanung in autonomen Systemen, wie Crane et al. [36]. Im Zusammenhang mit Fahrzeugen verwendet Reichardt [120] in einem seiner verfolgten Ansätze Gefahrenkarten, um das Fahrzeug autonom fahren zu lassen (siehe Abbildung 2-9.b). Dabei sind nicht nur Hindernisse auf der Karte in Form von Gefahrenpotentialen abgebildet, sondern auch andere Merkmale, wie Spurhaltung, Geschwindigkeitsbegrenzungen, andere Verkehrsregeln





**Abbildung 2-9:** Ansätze zur Gefahrenerkennung. a) Monte-Carlo-Simulation von Eidehall [47]; b) Gefahrenkarte von Reichardt [120]; c) Trajektorien-Cluster von Hervieu et al. [72] und d) Gefahrenkarten von Meysel [108].

und Fahrerintentionen. Jedes dieser Merkmale wird auf einer eigenen vorläufigen Karte dargestellt und die gesamte Karte wird dann aus den Einzelnen zusammengesetzt (aufsummiert oder gewichtet, u.a.). Die Vorteile des Ansatzes sind vor allem seine Modularität und die kontinuierliche Modellierung der Umgebung. Trotzdem sieht der Autor darin einige kritische Punkte, wie den hohen Rechenaufwand, die Mehrdeutigkeit in der Erstellung und dem Zusammenbringen der Karten, die mangelnde Nachvollziehbarkeit der erkannten Gefahren und die schwierige Modellierung der Fahrscenendynamik.

Weitere auf Gefahrenkarten basierte Ansätze sind zur Spurführung (Gerdes, Rossetter et al. [62, 63]), zur Bahnplanung (Sattel und Brandt [127] sowie Form et al. [52]), und zur Unterstützung des Fahrers (Holzmann [74]) vorgestellt.

Auch auf die Ursache-Wirkung fokussiert ist die Nutzung eines komplexen Regelsatzes. Reichardt [120] schlägt einen solchen Ansatz als Alternative zur Lösung der Probleme mit der Gefahrenkarte vor. Darin sind mehrere Merkmale der Fahrssituation in Zusammenhang gebracht. Inspiriert durch die Arbeiten von Reichardt berichten Marstaller und Bubb in [104] über das Konzept der Risikobewertung. Zur Ermittlung von kritischen Fahrssituationen müssen alle denkbaren Fehlerquellen identifiziert werden, so dass ein Assistenz- oder Automationssystem das Risiko effektiv reduzieren kann.

### 2.5.2 Gefahren als atypische Situationen

Unter dem Blickwinkel von Gefahr als statistisch seltene Situation sind in der Literatur Systeme aus dem Überwachungsbereich zu finden. Dabei sind üblicherweise die zu suchenden Ereignisse nicht direkt festgelegt, sondern durch Abweichung von der Normalität gegeben. Diese muss *a priori* gegeben oder angelernt sein. Darin basiert eine erste Gruppe von Ansätzen auf Trajektorieninformationen über die Objekte in der Umgebung. Die trajektorienbasierte Gefahrenerkennung erfolgt meist anhand von Daten statischer Plattformen. Dabei werden Merkmale wie Richtung und Krümmung aus den Trajektorien extrahiert, um diese zu klassifizieren (Clustern).

In vielen Arbeiten werden Cluster in Bildkoordinaten erzeugt, wie von Hervieu et al. [72] (siehe Abbildung 2-9.c) und Fu et al. [57]. Trajektorien, die zu keinem Cluster passen, werden als atypisch eingestuft. Im Gegenteil dazu erzeugen Meysel, Reulke et al. [108, 121] (Abbildung 2-9.d) statistische Karten in Weltkoordinaten anhand mehrerer Sensoren. Bewegt sich ein Objekt an einer Kartenposition anders als in der Karte gespeichert, so wird dies als atypisches Ereignis klassifiziert. Somit ist der Ansatz unabhängig von Trajektorien, Signaturen der gesuchten Ereignisse und semantischen Informationen über die Szene.

Eine weitere Gruppe benutzt statt Trajektorien andere Merkmale der Umgebung, um atypische Ereignisse zu erkennen. Zum Beispiel verwenden Utasi und Czúni [148] den optischen Fluss in Bildern. Die Arbeiten von Cenedese et al. [31] basieren auf normalisierten Bildkoordinaten für eine statistische Karte.

Alle diese Methoden benötigen statistische Kenntnisse über den normalen Raum (Karte oder Cluster), die sie über Training oder *a priori* Wissen gewinnen können. Zusätzlich muss die Bedeutung eines atypischen Ereignisses in weiteren Stufen eingeschätzt werden (*danger assessment*), denn nicht alle atypischen Ereignisse sind gleichzusetzen mit einer Gefahr. Der Nachteil ist, dass das Umgekehrte auch stimmt: nicht alle Gefahren stellen atypische Ereignisse dar.

## 2.6 Zusammenfassung des Kapitels

In diesem Kapitel wurde der Stand der Technik zur Objekt- und Gefahrenerkennung aufgearbeitet.

Für die Objekterkennung wurden unterschiedliche Sensoren vorgestellt, und deren Stärken und Schwächen tabellarisch zusammengefasst. Die üblichen Sensoren im Automotive-Bereich sind Radarsysteme, Lidars, Laserscanner und Kameras. Während Radarsensoren, Lidars und Laserscanner eine höhere Zuverlässigkeit als Vision-Systeme aufweisen, haben die letztgenannten eine höhere Winkelauflösung und können im Bereich der Klassifikation besser eingesetzt werden.

Für die Fusion der Sensordaten sind unterschiedliche Abstraktionsebenen möglich. Je sensornäher die Fusion geschieht, desto performanter kann sie implementiert werden. Allerdings weisen solche Systeme auch eine geringere Flexibilität und Erweiterbarkeit auf. Im Gegenteil dazu können Ansätze zur Objekt- oder Tracking-Fusion allgemeiner und skalierbarer, jedoch in der Regel weniger performant und genau implementiert werden.

Bezüglich der Gefahrenerkennung wurden zwei Gruppen von Ansätzen vorgestellt. Die Erste ba-

siert auf der Ursache-Wirkungs-Beziehung zwischen Gefahr und Unfall. Darin sind die meisten herkömmlichen Assistenz- und Automationssysteme sowie weitere Systeme eingeschlossen, welche die Gesamtsituation in Form von Gefahrenkarten oder Regelsätzen betrachten. Die zweite Gruppe beschäftigt sich mit statistisch atypischen Ereignissen, die eventuell Gefahren darstellen können.



### 3 Auswahl und Definition des Fusionsansatzes und des Objektmodells

Dieses kurze Kapitel stellt den Rahmen der Arbeit zur Objekterkennung dar und fungiert als Bindeglied zu den darauf folgenden Kapiteln. Im ersten Abschnitt wird über die möglichen Fusionssebenen und Paradigmen diskutiert. Zudem wird gezeigt, dass die Fusion auf der bevorzugten Objektebene und auf der Merkmalsebene mathematisch äquivalent ist. Da die Objektebene jedoch einige wichtige Vorteile gegenüber der Merkmalsebene besitzt, wird diese für die restliche Arbeit ausgewählt. Für die Fusion auf Objektebene ist es notwendig, dass Objektbeobachtungen mit Genauigkeits- und Zuverlässigkeitsinformationen in fahrzeugzentrierten Koordinaten aus jedem einzelnen Sensor erzeugt werden (siehe folgende Kapitel). Sowohl die aus den einzelnen Sensoren extrahierten Objektbeobachtungen als auch die verfolgten und fusionierten Objekthypothesen werden anhand desselben Modells dargestellt. Die Grundsteine dieses Modells, das für den weiteren Verlauf der Arbeit verwendet wird, werden im Abschnitt 3.2 definiert.

#### 3.1 Fusionsansatz

Es wurde für diese Arbeit ein konkurrierender Ansatz zur Objekterkennung auf der Objektebene ausgewählt. Auf dieser Ebene ist die Abstraktion von den Sensoren höher als bei einer sensor-nahen Fusion, womit hohe Verfügbarkeit (bei Sensorausfall) und Erweiterbarkeit gewährleistet werden. Im Vergleich mit einer höheren Ebene (wie Tracking-Fusion) können mit diesem Ansatz Datenfehler und Ausreißer besser behoben werden, womit eine höhere Zuverlässigkeit gewährleistet wird.

Um die eingehenden Sensordaten in das Sensorfusionssystem durch die Assoziation einzubringen, müssen diese und die bereits im System vorhandenen Objekthypothesen in einem gemeinsamen Koordinatensystem dargestellt werden (*alignment*). Auf der Objektebene (auf Basis der Begriffsdefinitionen in 2.2 und 2.4.2) ergeben sich zwei Möglichkeiten: die Merkmalsfusion und die eigentliche Objektfusion. Diese werden im Folgenden diskutiert.

Zum Vergleich von Merkmals- und Objektfusion wird hier die Beobachtungsgleichung des Kalman-Filters herangezogen (siehe Details in Anhang C). Der Zusammenhang zwischen den Objekthypothesen mit Zustandsvektor  $\mathbf{x}$  und den Beobachtungen mit Messvektor  $\mathbf{z}$  ist im Zeitpunkt  $t_k$  durch die Beobachtungsmatrix  $\mathbf{H}$  und das Messrauschen  $\mathbf{v}$  (weißes Rauschen mit Kovarianzmatrix  $\mathbf{R}$ ) gegeben:

$$\mathbf{z}(t_k) = \mathbf{H}(t_k) \cdot \mathbf{x}(t_k) + \mathbf{v}(t_k) \quad (3.1)$$

Eine Fusion auf Merkmalsebene bedeutet, dass die Assoziation auf Sensordatenmerkmalen stattfindet, wie beispielsweise Reflexionspunkten für den Laserscanner und Kanten oder Eckpunkten in Bildern. Hierdurch wird die Sensordatenverarbeitung einfacher, die Abbildung der prädizierten

Objekthypothesen in den Merkmalsraum ist jedoch komplexer. Letzteres setzt genaue Kenntnis der Objekte und der Sensoren voraus. Die Verarbeitung von nichtassozierten Merkmalen ist auch komplexer, um neue Objekte im System zu initialisieren. Somit entsprechen der Messvektor, das Messrauschen und die Beobachtungsmatrix folgendem (hier mit Index  $m$  wie Merkmal):

$$\begin{aligned} \mathbf{z}_m(t_k) &= \text{Sensordatenmerkmal} \\ \mathbf{R}_m(t_k) &= \text{Merkmalsunsicherheit} \\ \mathbf{H}_m(t_k) &= \text{Transformation von Objekt zu Merkmal} \end{aligned}$$

$$\mathbf{z}_m, \mathbf{R}_m \rightarrow \boxed{\begin{array}{c} \mathbf{F}, \mathbf{Q}, \mathbf{H}_m \\ \text{Filter} \end{array}} \rightarrow \mathbf{x}, \mathbf{P}$$

Alternativ kann die Fusion auf Objektebene stattfinden. Dies bedeutet, dass die eingehenden Daten Objektbeobachtungen sind, was eine hohe Komplexität für die Sensordatenverarbeitung, hier als  $\mathbf{H}_t$  ausgedrückt, voraussetzt. Da die prädizierten Objekthypothesen bereits im selben Koordinatensystem sind wie die eingehenden Objektbeobachtungen, ist die Beobachtungsmatrix  $\mathbf{H}_o$  trivial (entspricht der Einheitsmatrix), womit der Kalman-Filter schlanker ist. Der Bezug auf die Merkmalsebene entspricht dabei  $\mathbf{H}_m = \mathbf{H}_t \cdot \mathbf{H}_o$ . Die Assoziation ist im Objektraum nachvollziehbarer und für alle Sensoren gleich. Die Initialisierung neuer Objekthypothesen aus Objektbeobachtungen ist ebenfalls direkter. Zusätzlich zu den erwähnten Vorteilen sind für die Fusion auf dieser Ebene keine sensorspezifischen Informationen notwendig, womit die Modularität und Erweiterbarkeit des Fusionssystems höher sind.

$$\begin{aligned} \mathbf{z}_o(t_k) &= \text{Objektbeobachtung} \\ \mathbf{R}_o(t_k) &= \text{Objektunsicherheit} \\ \mathbf{H}_o(t_k) &= \text{Identitätsmatrix bzw. Mapping} \end{aligned}$$

$$\mathbf{z}_m, \mathbf{R}_m \rightarrow \boxed{\begin{array}{c} \mathbf{H}_t \\ \text{Vorverarb.} \end{array}} \rightarrow \mathbf{z}_o, \mathbf{R}_o \rightarrow \boxed{\begin{array}{c} \mathbf{F}, \mathbf{Q}, \mathbf{H}_o \\ \text{Filter} \end{array}} \rightarrow \mathbf{x}, \mathbf{P}$$

Für den Kalman-Filter<sup>1</sup> kann bewiesen werden, dass die Filterung auf beiden Fusionssebenen mathematisch äquivalent ist, solange die Beobachtungen  $\mathbf{z}$ , die Messunsicherheiten  $\mathbf{R}$  sowie die Beobachtungsmatrix  $\mathbf{H}$  anhand der Varianzfortpflanzung (näher erläutert in Kapitel 5.5) entsprechend angepasst sind. Außerdem muss die sensorspezifische Transformation  $\mathbf{H}_t$  invertierbar sein, da aus den Sensordaten Objektbeobachtungen extrahiert werden müssen – dies entspricht der eingangs erwähnten Sensordatenverarbeitung. Eine detaillierte Herleitung wird im Abschnitt C.4 des Anhangs durchgeführt. Die Ausdrücke auf der linken Seite entsprechen dem Kalman-

---

<sup>1</sup>und damit auch für den Informationsfilter, da diese mathematisch äquivalent sind.

Filter auf Merkmalsebene, während auf der rechten Seite die angepassten Ausdrücke für die Objektebene stehen:

Merkmalsebene	Objektebene
$\mathbf{x}(t_k), \mathbf{P}(t_k)$	$\mathbf{x}(t_k), \mathbf{P}(t_k)$
$\mathbf{F}(t_k)$	$\mathbf{F}(t_k)$
$\mathbf{Q}(t_k)$	$\mathbf{Q}(t_k)$
$\mathbf{z}_m(t_k)$	$\mathbf{z}_o(t_k) \triangleq \mathbf{H}_t^{-1}(t_k)\mathbf{z}_m(t_k)$
$\mathbf{R}_m(t_k)$	$\mathbf{R}_o(t_k) \triangleq \mathbf{H}_t^{-1}(t_k)\mathbf{R}_m(t_k)(\mathbf{H}_t^{-1}(t_k))^T$
$\mathbf{H}_m(t_k) \triangleq \mathbf{H}_t(t_k)\mathbf{H}_o(t_k)$	$\mathbf{H}_o(t_k) \triangleq \mathbf{H}_t^{-1}(t_k)\mathbf{H}_m(t_k)$

Es sei angemerkt, dass sich die Transformationen  $\mathbf{H}_t$  und  $\mathbf{H}_t^{-1}$  in der Regel nicht als einzelne Matrizen ausdrücken lassen, sondern eine gewisse (nichtlineare) Logik bzw. Algorithmik einschließen. Diese Matrizen tauchen im Filter auf Objektebene nicht auf, da sie bereits in der Vorverarbeitung eingesetzt wurden.

Aufgrund der hier erwähnten Vorteile wird die Fusion in dieser Arbeit auf Objektebene durchgeführt. Die Entwicklung des hierfür verwendeten Objektmodells wird nun im folgenden Abschnitt näher erläutert.

## 3.2 Modell für Objektbeobachtungen und Objekthypothesen

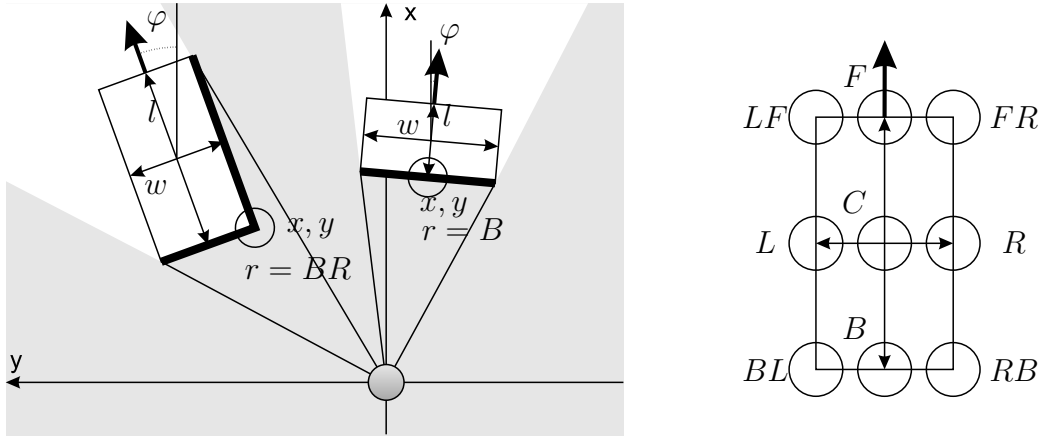
In diesem Abschnitt wird das Modell eingeführt, mit dem im Rahmen dieser Arbeit alle erkennbaren Objekte dargestellt werden. Da die Fusion auf Objektebene stattfindet, gilt das Modell sowohl für die Objektbeobachtungen aus der Sensordatenverarbeitung als auch für die gefilterten Objekthypothesen in der Verfolgungsphase.

Die in der Literatur verwendeten Modelle sind sehr unterschiedlich (siehe Abschnitte 2.3 und 2.4), wie Punktmodelle, Zwei-Linien-Modelle und Quadermodelle. Es wird hier ein allgemeines Modell angestrebt, mit dem ausgedehnte Objekte dargestellt werden können und das jedoch minimal ist. Ein ausgerichtetes Quadermodell ist dafür geeignet – siehe linke Seite der Abbildung 3-1.

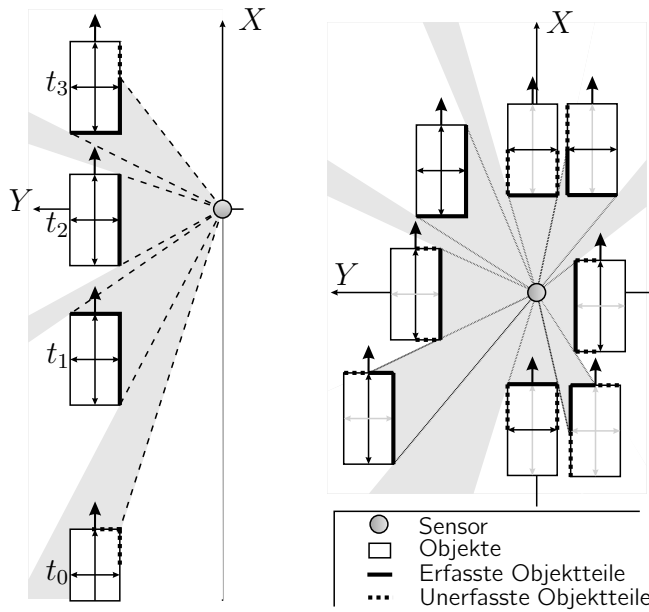
Das Objektmodell schließt seine Position  $(x, y)$ , Ausdehnung  $(w, l)^2$ , Ausrichtung  $(\varphi)$  und relative Geschwindigkeit<sup>3</sup>  $(\dot{x}, \dot{y})$  ein. Alle Koordinaten sind in fahrzeugzentrierten Koordinaten angegeben. Da keine Höhen berücksichtigt werden, entsprechen diese Koordinaten der Straßenebene und werden an manchen Stellen dieser Arbeit mit Index  $r$  verwendet, um sie von Bildkoordinaten oder globalen Koordinaten zu unterscheiden. Hierbei entsprechen die  $X$ -Achse der Fahrzeuglängsrichtung und die  $Y$ -Achse der Querrichtung nach links. Da es für die Sensoren häufig nicht möglich ist, die Seiten eines Objektes zu unterscheiden (vorn, hinten, rechts und links), wird hier die Ausrichtung immer um die  $0^\circ$  festgelegt, womit  $\varphi \in (-45^\circ, 45^\circ]$ . Damit ist die Länge immer die Ausdehnung in  $X$ -Richtung und die Breite die in  $Y$ -Richtung. Der Zustandsvektor ergibt sich aus den aufgeführten Variablen wie folgt:

<sup>2</sup>Die Objekthöhe wird im Rahmen der Arbeit nicht berücksichtigt.

<sup>3</sup>Auch die Geschwindigkeit ist enthalten, wird jedoch erst in späteren Abschnitten, mit der Verfolgung, behandelt.



**Abbildung 3-1:** Verwendetes ausgerichtetes Quadermodell (rechts) und mögliche Referenzen eines Objektes, dargestellt durch Kreise (links).



**Abbildung 3-2:** Verändern des Aussehens eines Objektes im Laufe der Zeit (links) und bei unterschiedlichen Anordnungen zu anderen Objekten und zum Sensor (rechts).

$$\mathbf{x}(t_k) = [x \ y \ \dot{x} \ \dot{y} \ \varphi \ w \ l]^T \quad (3.2)$$

Zu diesem Zustandsvektor gehört eine Kovarianzmatrix  $\mathbf{P}$ , die Aufschluss über die Datenunsicherheiten gibt. Für Objektbeobachtungen werden die Bezeichnungen Beobachtungsvektor  $\mathbf{z}$  und Messunsicherheit  $\mathbf{R}$  mit demselben Modell verwendet.

Ein besonderes Problem in der Objekterkennung stellt die veränderliche Beobachtbarkeit der Objekte dar. Bedingt durch die Beobachtungsgeometrie zwischen Sensorplattform und den umgebenden Objekten sowie durch Teilverdeckungen zwischen den Objekten ändert sich die Beobachtbarkeit der unterschiedlichen Objektkomponenten, wie in Abbildung 3-2 verdeutlicht. Hier zeigt sich einer der Vorteile eines Fusionssystems, da damit der komplette Objektzustand im Laufe der Zeit geschätzt werden kann. Es ist jedoch nicht möglich, alle Objektbeobachtungen und



Objekthypothesen anhand einer festen Referenz, wie beispielsweise Objektmittelpunkt, Schwerpunkt oder Heckmittelpunkt, zu beschreiben, denn dies würde die Genauigkeit der Schätzung erheblich beeinträchtigen.

Aus diesem Grund und um Redundanzen zu vermeiden, wird das Objekt stets anhand der Position eines seiner Punkte dargestellt. Der Punkt, auf den sich die Position bezieht, wird in Form einer logischen Referenz ( $r$ ) ausgedrückt, die jede der Objektecken, Mitte der Objektseiten sowie Objektmittelpunkt betragen kann:  $r \in \{C$  (*center*, Objektmitte),  $B$  (*back*, Heckmittelpunkt),  $BL$  (*back-left*, hintere linke Ecke),  $L$  (*left*, Mitte der linken Seite),  $LF$  (*left-front*, linke vordere Ecke),  $F$  (*front*, Mitte der vorderen Seite),  $FR$  (*front-right*, vordere rechte Ecke),  $R$  (*right*, Mitte der rechten Seite) oder  $RB$  (*right-back*, rechte hintere Ecke) $\}$  – siehe rechte Seite der Abbildung 3-1. Die logische Referenz wird variabel gestaltet, so dass diese je nach Situation angepasst werden kann.

Zusätzlich wird in der Verarbeitung der Sensordaten und der Verfolgung bzw. Fusion besonderer Wert auf die partielle Beobachtbarkeit der einzelnen Komponenten des Objektmodells gelegt. Analog zu dem von Stüker [144] eingeführten Initialisierungsvektor, begleitet hier jedes Objekt eine zusätzliche Komponente, die Beobachtbarkeitsmatrix  $\mathbf{V}_x$  (bzw.  $\mathbf{V}_z$ ):

$$\mathbf{V}_x(t_k) = \begin{bmatrix} v_{x11} & 0 & \cdots & 0 \\ 0 & v_{x22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & v_{x_{n_x n_x}} \end{bmatrix}; v_{x_{ii}} = \begin{cases} 1 & \text{wenn } x_i \text{ beobachtbar,} \\ 0 & \text{sonst.} \end{cases} \quad (3.3)$$

Zusammenfassend besteht das Objektmodell aus einem Zustandsvektor  $\mathbf{x}$  bzw. Messvektor  $\mathbf{z}$ , einer Kovarianzmatrix  $\mathbf{P}$  bzw. Messunsicherheitsmatrix  $\mathbf{R}$ , einer logischen Objektreferenz  $r_x$  bzw.  $r_z$  und einer Beobachtbarkeitsmatrix  $\mathbf{V}_x$  bzw.  $\mathbf{V}_z$ . Auf diesen Strukturen basieren die Objekterkennungsmethoden, welche in den nächsten Abschnitten vorgestellt werden.

Das in dieser Arbeit verwendete Dynamikmodell entspricht einer geradlinigen Bewegung mit konstanter Geschwindigkeit, wie im Zustandsvektor (3.2) ersichtlich ist. Komplexe Modelle haben viele Freiheitsgrade zufolge. Damit sind viele genaue Beobachtungen notwendig, um den Objektzustand zu schätzen (vgl. Cramer [35]). Einfache Modelle, wie das in dieser Arbeit verwendete, lassen sich dagegen schneller initialisieren und sind wegen besserer Konvergenz stabiler. Da es keine nichtlinearen Anteile besitzt, benötigt die Filterung keine Näherung, ist also optimal. Ein vereinfachtes Dynamikmodell besitzt jedoch auch Nachteile. Die Bewegung von manövrierenden Objekten lässt sich damit nicht beschreiben (wie beim Beschleunigen oder in Kurven).

Da das Modell relative Positionen verwendet, sind die Objektbewegungen stets denen des eigenen Fahrzeugs überlagert. Diese Effekte bewirken Schwierigkeiten bei der Assoziation und Stabilität des geschätzten Objektzustands.

Um diese Effekte zu beseitigen, werden häufig in der Literatur adaptive Verfahren vorgeschlagen, siehe Bar-Shalom [10, 11]. Beispielsweise wird das Prozessrauschen mit einem zusätzlichen Faktor belastet (*fudge factor*) oder es werden zusätzliche Komponenten im Zustandsvektor hin-

zugefügt oder entfernt. Ein weiteres Verfahren ist das *Interacting Multi Model Filter* (IMM), bei dem mehrere Filter mit unterschiedlichen Dynamikmodellen verfolgt werden und immer nur der passende eingeschaltet wird. Diese Ansätze gehen von einer fehlerfreien Assoziation aus. Die Überlagerung der Objektbewegung mit der eigenen Bewegung ist nach [144] möglich, bringt jedoch Nichtlinearitäten in das Modell ein. Daher wurde im Rahmen dieser Arbeit darauf verzichtet.

In dieser Arbeit werden unterschiedliche Strategien verwendet, um mit den erwähnten Herausforderungen umzugehen. Auf der einen Seite werden im Objektmodell die erwähnten Strategien der variablen, logischen Referenz  $r$  sowie die partielle Beobachtbarkeit der Objekthypothesen  $\mathbf{V}$  definiert. Dies ermöglicht eine konsistente Verfolgung der Objekte, auch bei Veränderungen des beobachtbaren Objektteils. Dadurch hat diese Fusionsstrategie auch einen komplementären Charakter. Auf der anderen Seite werden in Kapitel 7 Strategien vorgestellt, um gegen Ausreißern, manövrierende Objekte, und *split*- und *merge*-Probleme entgegenzuwirken, wie beispielsweise die mehrfache Assoziation, die Verdoppelungs- sowie die Vereinigungsstrategie. Im Gegenteil zur IMM sind diese Strategien schneller und benötigen keine Annahmen über mögliche Objektbewegungen.

## 3.3 Zusammenfassung des Kapitels

In diesem Kapitel wurden die grundsätzlichen Ansätze zur Fusionierung der Sensordaten (Fusionsebenen) vorgestellt und diskutiert. Dabei wird der für diese Arbeit ausgewählte Ansatz auf Objektebene genauer mit der Merkmalsebene verglichen. Beide Ebenen sind mathematisch äquivalent, sofern Objektbeobachtungen aus den einzelnen Sensoren extrahiert und deren Messunsicherheiten angepasst werden. Hierbei besitzt die Objektebene wichtige Vorteile, wie Modularität, Flexibilität und einfachere Erweiterbarkeit. Außerdem sind dabei die Assoziation und die Initialisierung direkter möglich und der Filter kann schlanker gestaltet werden.

Ausgehend von der Objektebene konnte im zweiten Abschnitt dieses Kapitels das Objektmodell ausgewählt und entwickelt werden. Es soll für alle extrahierten Objektbeobachtungen sowie für die verfolgten bzw. fusionierten Objekthypothesen verwendet werden. Darin sind wichtige neue Aspekte definiert worden, wie die variable, logische Referenz und die partielle Beobachtbarkeit der Objekte. Damit beinhaltet dieses Modell detaillierte Informationen in einer kompakten Form.

## 4 Verarbeitung auf Sensorebene

Dieses Kapitel beschreibt die notwendigen Verarbeitungsschritte, die vor der eigentlichen Objekterkennung notwendig sind (siehe Zusammenhang mit der gesamten Arbeit in Abbildung 4-1). Diese umfassen die Merkmalsextraktion aus den Sensordaten, die Korrektur von Fahrzeugschwankungen während der Fahrt und die Ermittlung der Fahrkorridorbegrenzungen.

Aus Kamerabildern werden zunächst Kanten-, Geraden- und Texturmerkmale extrahiert. Die Laserscannerdaten werden im Sensor vorverarbeitet und segmentiert.

Um Fahrzeugschwankungen während der Fahrt in den Sensordaten zu kompensieren, wird hier eine Bildverarbeitungsmethode entworfen, mit der die Ausrichtung der Kamera (und damit die der gesamten Sensorplattform) bezüglich der Straßenkoordinaten bestimmt wird. Die Methode basiert auf den geometrischen Eigenschaften der Fahrstreifenmarkierungen. Werden diese aus dem Bild extrahiert, so lässt sich damit über eine Parameterausgleichung auf die Kameraorientierung schließen.

Zusätzlich wird in diesem Kapitel ein Verfahren zur Erkennung des Fahrkorridors vorgestellt, auf den die Aufmerksamkeit der weiteren Verarbeitungsschritte zur Objekterkennung konzentriert wird. Dies ermöglicht eine Reduktion von seitlichem Rauschen und eine effiziente Verarbeitung. Die verwendeten Daten sind Kamerabilder (Fahrstreifendaten), Informationen über die befahrene Straße aus einer digitalen Karte sowie Laserscannerdaten. Da keiner der drei Sensoren eine ausreichend hohe Zuverlässigkeit in allen Fahrsituationen gewährleistet, wird eine sensornah, kooperative Fusion entwickelt. Jeder Sensor schränkt dabei den Suchbereich ein, in dem die Fahrkorridorbegrenzungen liegen können.

### 4.1 Verarbeitung der Kamerabilder

In diesem Abschnitt werden die Schritte der Datenvorverarbeitung, Extraktion von Grundmerkmalen, Binarisierung und Teile der Segmentierung für die Kamerabilder vorgestellt. Die in dieser Arbeit benötigten Grundmerkmale sind Kanten, Geraden und Textur im Bild.

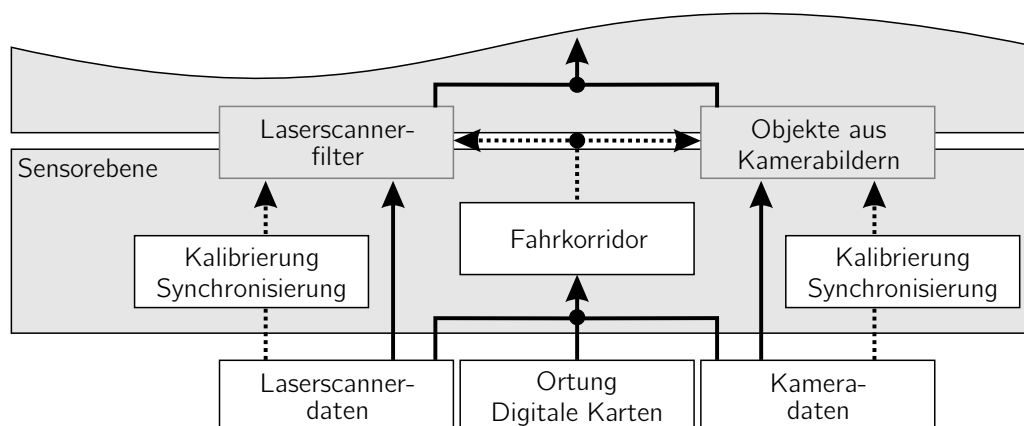
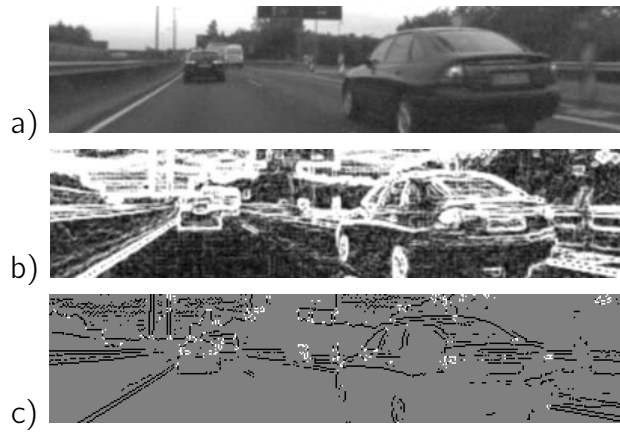


Abbildung 4-1: Verarbeitung auf Sensorebene im Rahmen der Arbeit.



**Abbildung 4-2:** Kantenfilterung: a) das Originalbild, b) das Gradientenbild und c) das Kanten- und Eckenbild (jeweils schwarz und weiß).

### 4.1.1 Kantenfilterung

Zur Bestimmung der Kanten und Eckpunkte im Bild wird das von Wohlfeil et al. in [165] vorgestellte Verfahren verwendet. Zuerst wird das Bild mit einem Optimum-Kantenfilter nach Ando [7] gefaltet. Daraus ergeben sich Gradienteninformationen pro Pixel in horizontaler und vertikaler Komponente ( $g_X$  und  $g_Y$ ), was auf die Kantenausrichtung schließen lässt. Zur Ausfilterung vom Rauschen im Kantenbild werden auch ein Schwellwert und eine *Non-maxima-suppression* verwendet, womit nur lokale Maxima beibehalten werden. Anschließend wird die Umgebung um jedes Pixel mit der Kovarianzmatrix der horizontalen und vertikalen Gradienten charakterisiert und eine Hauptachsentransformation entsprechend Kanade und Tomasi [146] durchgeführt. Die erhaltenen Eigenwerte geben Aufschluss über das Pixel. Sind sie klein, so handelt es sich um einen homogenen Bereich. Sind sie groß und weisen eine klare Richtung auf, so handelt es sich um einen Kantenpunkt. Weisen sie unterschiedliche Richtungen auf, so handelt es sich um einen Eckpunkt.

Die Vorteile dieses Verfahrens sind vor allem dessen niedriger Rechenaufwand und die Robustheit gegenüber Rauschen und Artefakten. Abbildung 4-2 zeigt ein Beispiel für ein Bild, das entsprechende Gradientenbild und die extrahierten Kanten und Eckpunkte.

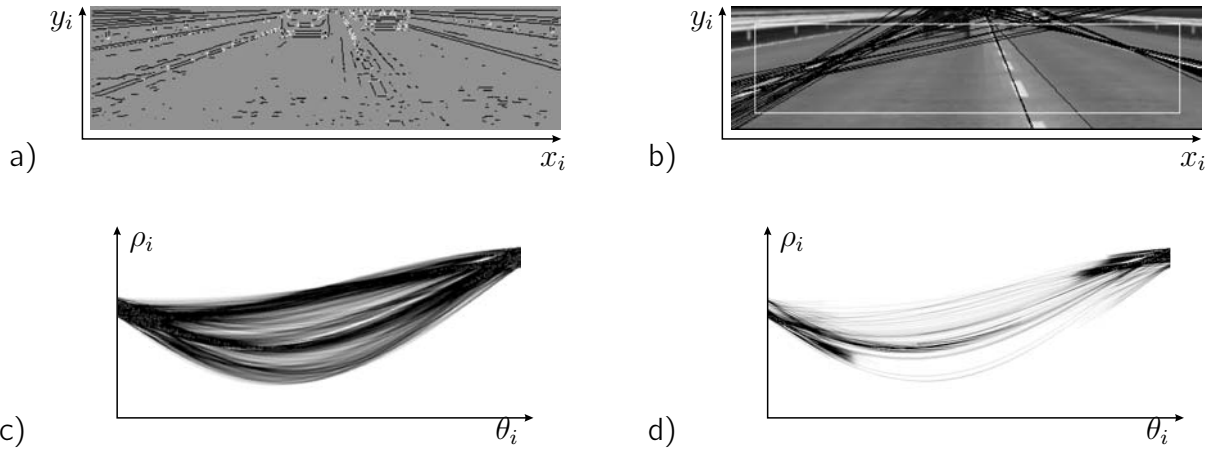
### 4.1.2 Extraktion von Geraden im Bild

Für die Extraktion von lückenhaften Geraden im Bild wird hier die Hough-Transformation verwendet. Dafür werden die gesuchten Strukturen (in diesem Fall Geraden) in einem akkumulierenden Parameterraum (Hough-Raum) dargestellt. Siehe Abbildung 4-3.c und d (dunklere Zellen haben höhere akkumulierte Werte). Sei eine Gerade im Bild wie folgt beschrieben<sup>1</sup>:

$$x_i \cdot (-\sin \theta_i) + y_i \cdot \cos \theta_i = \rho_i \quad (4.1)$$

Jeder Bildpunkt  $[x_i \ y_i]^T$  entspricht einer Menge von Geraden  $[\theta_i \ \rho_i]^T$  im Hough-Parame-

<sup>1</sup>Bildkoordinaten mit Ursprung unten links.



**Abbildung 4-3:** Beispiel der Hough-Transformation. a) Kantenbild (Bildraum), c) Hough-Parameter-raum (mit großem  $p_{H\theta}$ ), d) optimierte Hough-Transformation (mit kleinem  $p_{H\theta}$ ) und b) erkannte Geraden (Bildraum).

terraum (einer Sinuskurve), die das Geradenbündel darstellt, das durch den Punkt läuft. Eine Gerade  $[\theta_i \ \rho_i]^T$  in Hough-Parameterraum entspricht wiederum allen Bildpunkten  $[x_i \ y_i]^T$ , die auf dieser Gerade liegen.

Die Methode erfolgt in zwei Schritten. Zuerst werden alle als Kantenpunkt klassifizierten Pixel in den Hough-Parameterraum transformiert und akkumuliert. Als Optimierung wird die Transformation um den jeweiligen Kantenwinkel

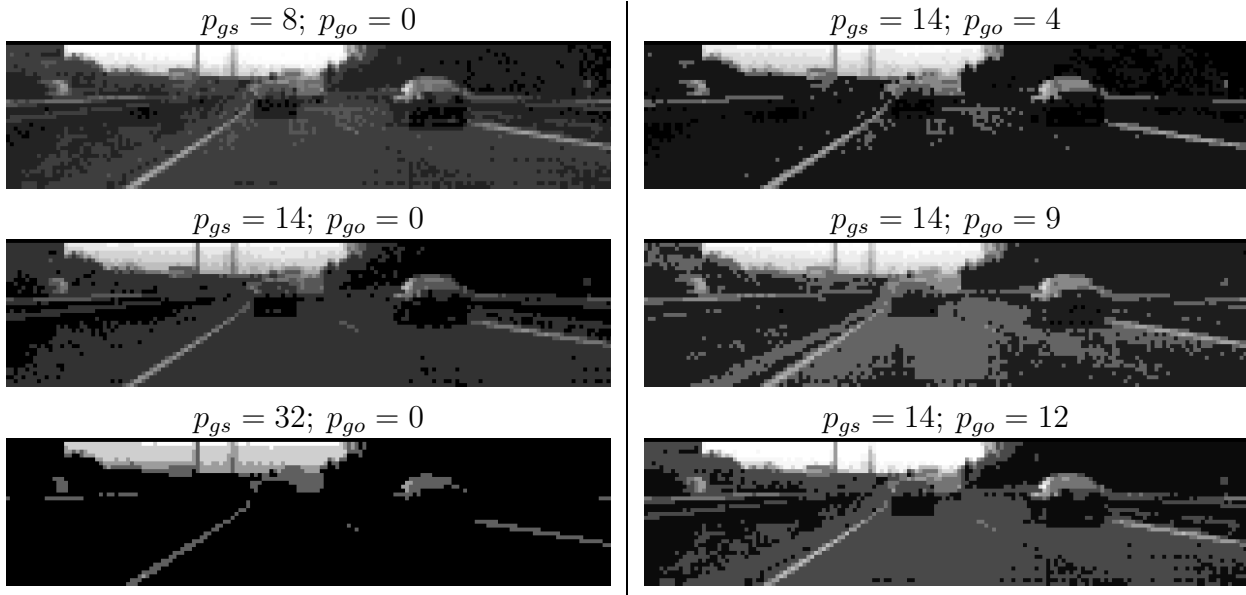
$$\theta = \arctan(-g_X/g_Y) \pm p_{H\theta}, \quad (4.2)$$

eingeschränkt, mit  $p_{H\theta}$  als einem kleinen fixen Winkelabstand. Weitere Optimierungen werden durch Einschränkung des Bildsuchraumes und Hough-Raumes erreicht. Im zweiten Schritt werden die lokalen Maxima aus dem Hough-Raum extrahiert und als Ergebnis ausgegeben.

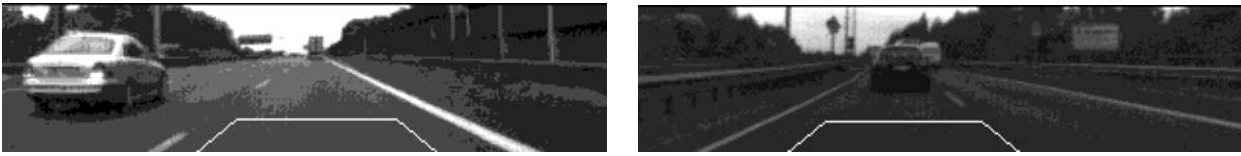
### 4.1.3 Berechnung der Bildtextur

Nach Lippok [93] lassen sich die Ansätze zur Berechnung von Bildtexturen wie folgt einteilen: statistische Merkmale (wie Energie, Kontrast, Mittelwert, Entropie oder Homogenität anhand der *Cooccurrence*-Matrix von Haralick [71]), modellbasierte Merkmale (wie Markov-Zufallsfelder oder Gibbs-Zufallsfelder), geometrische Merkmale und andere. Diese Ansätze sind sehr rechenaufwändig und liefern keine allgemeine, belichtungs- und skalierungsunabhängige Lösung. Deswegen sind sie für perspektivische Bilder aus dem Fahrzeug heraus nicht geeignet.

Als Textur wird hier der lokale Grauwert um jedes Pixel verwendet, welcher in festen Grauwertstufen quantifiziert wird. Dieser ist leicht zu berechnen und liefert ausreichend gute Segmentierungsmöglichkeiten. Die Berechnung der Grauwerttextur im Bild hängt von zwei Parametern ab, der Grauwertstufe (*grey step*,  $p_{gs}$ ) und der Verschiebungskonstante (*grey offset*,  $p_{go}$ ). Ersterer bestimmt die Anzahl von Stufen oder Kontrast im Texturbild und damit auch die Anzahl von Segmenten mit homogener Grauwerttextur. Letzterer bestimmt die Grenzen der Stufen und verschiebt damit auch unklare Segmentbegrenzungen im Bild bei gleicher Anzahl von Stufen



**Abbildung 4-4:** Effekt von Grauwertstufe  $p_{gs}$  und Verschiebungskonstante  $p_{go}$  in einem Texturbild-ausschnitt. Links wird die Grauwertstufe und rechts die Verschiebungskonstante geändert.



**Abbildung 4-5:** Texturbild in zwei unterschiedlich beleuchteten Situationen. Der Referenzbereich ist weiß markiert.

oder Kontrast. Abbildung 4-4 verdeutlicht die Effekte der Änderung dieser Parameter auf einem Beispielbild.

Die Textur der Bilder variiert sehr stark je nach Situation und Beleuchtung. Eine robuste auf Textur basierte Methode muss deshalb adaptiv gestaltet werden. Ähnlich wie Ulrich et al. in [147] vorschlagen, wird hier angenommen, dass sich keine Objekte oder Markierungen unmittelbar vor dem Fahrzeug befinden, womit die entsprechende Bildregion als Straße definiert ist (siehe Abbildung 4-5). Sei die Grauwertspanne in der markierten Bildregion durch  $g_{max}$  und  $g_{min}$  gegeben, so lassen sich  $p_{gs}$  und  $p_{go}$  wie folgt errechnen:

$$p_{gs} = g_{max} - g_{min} + 1 \quad (4.3)$$

$$p_{go} = p_{gs} - (g_{min} \bmod p_{gs}) \quad (4.4)$$

Befinden sich doch Straßenmarkierungen im Referenzbereich, so wird dessen Grauwertspanne fälschlicherweise sehr hoch eingeschätzt, womit die Grauwerttextur des Bildes auch verfälscht ist (Ergebnis ähnlich wie in Abbildung 4-4 unten links). Um dies zu vermeiden, werden die Werte  $g_{max}$  und  $g_{min}$  genommen, bei denen sich ein Sprung im Histogramm zeigt. Wenn dieser nicht deutlich ist, wird das obere bzw. untere Quantil genommen.

## 4.2 Aufbereitung der Laserscannerdaten

Der für diese Arbeit verwendete Laserscanner verfügt über eine eigene Vorverarbeitung, welche die Korrektur und Objektextraktion (Segmentierung der Laserscannerreflexionen in Objekten) sowie deren Verfolgung (Tracking) durchführt (vgl. [59]). Zu den durchgeführten Schritten gehören die Erkennung von Schmutz, Regen und Bodenreflexionen sowie eine Korrektur der Fahrzeugbewegung während der Aufnahme. Zur Segmentierung der Scanpunkte in Konturen wird deren entfernungsabhängiger Abstand betrachtet. Des Weiteren werden die Konturen in der Verarbeitungseinheit getrackt und klassifiziert, was in dieser Arbeit nicht weiter verwendet wird, da eine eigene Objektverfolgung entwickelt wird.

## 4.3 Korrektur der Fahrzeugschwankungen

Um den Bezug von den Sensordaten zum übergeordneten Fahrzeugkoordinatensystem<sup>2</sup> zu gewinnen, werden die Alignment-Parameter verwendet, d.h. für jeden Sensor drei Translationsparameter  $t_x$ ,  $t_y$  und  $t_z$  und drei Rotationsparameter  $\omega$ ,  $\phi$  und  $\kappa$  (in die jeweilige Raumrichtung, in dieser Reihenfolge). Es wird hier angenommen, dass sich diese während der Fahrt nicht ändern, da Kamera und Laserscanner fixiert sind. Um die Koordinaten auf die Straßenebene beziehen zu können, ist außerdem die Kenntnis der Ausrichtung des Fahrzeugs zur Straße notwendig. Diese verändert sich geringfügig während der Fahrt, bedingt durch Unebenheiten der Straße, Beschleunigungs- und Bremsvorgänge sowie Kurvenfahrten. Für diese Arbeit wird die Ausrichtung der Kamera zur Straßenebene untersucht und auf die Trägerplattform auf Basis der Alignment-Parameter übertragen. Solange die Schwankungen klein sind, kann die auf dem Bild erkannte Orientierungsänderung näherungsweise auf das Fahrzeugkoordinatensystem übertragen werden. Die Vernachlässigung dieser Abweichungen hätte negative Folgen auf die weitere Verarbeitung der Sensordaten, wie Ungenauigkeit oder Unzuverlässigkeit, [49]. Die hier vorgestellte bildverarbeitungs-basierte Methode wurde zum Teil vom Autor bereits in [30] vorgestellt.

Zur Bestimmung der äußeren Orientierung werden in der Photogrammetrie Markierungen oder Passpunkte verwendet, deren Positionen in Objekt- und Bildraum vorher genau bestimmt wurden (vgl. [96]). Aus den Bildkoordinaten einer genügenden Anzahl von Passpunkten lassen sich die äußeren Orientierungsparameter bestimmen (räumlicher Rückwärtsschnitt, DLT<sup>3</sup>, Bündelausgleich etc. vgl. [96]). Andere Ansätze verwenden deutlich verfolgbare Features im Bild, um in einer Sequenz die Veränderung der äußeren Orientierung einzuschätzen, wie Maybank et al. [105], Luong et al. [97] und Dron [44]. In den Arbeiten von Wohlfeil et al. [165], Krüger [87] und Rasmussen [118] werden alternative Methoden vorgestellt, die den Fluchtpunkt im Bild bestimmen und dadurch die Orientierung in  $Y$ - und  $Z$ -Richtung (sprich  $\phi$  Nick- und  $\kappa$  Gierwinkel) kompensieren. Eine weitere Möglichkeit, die Fahrzeugschwankungen auszugleichen, besteht in der Verwendung von zusätzlichen Beschleunigungssensoren<sup>4</sup>. In diesem Fall ist eine hochgenaue relative Orientierung und Synchronisierung der Systeme notwendig.

In dieser Arbeit wird eine Methode vorgestellt, die auf den geometrischen Eigenschaften der

<sup>2</sup>X-Komponente in Fahrzeuglängsrichtung, Y-Komponente in Querrichtung nach links.

<sup>3</sup>DLT – *Direct Linear Transformation*

<sup>4</sup>Zum Beispiel ein IMU-System (*Inertial Measurement Unit*)

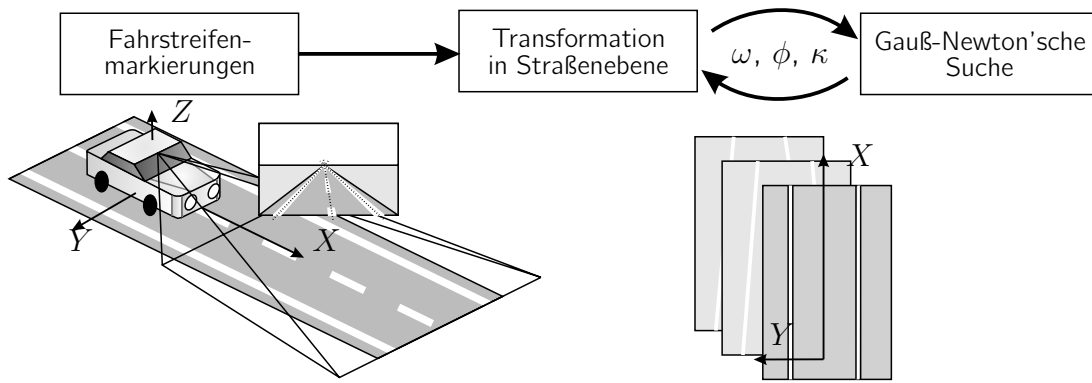


Abbildung 4-6: Schritte der Bestimmung der Fahrzeugorientierung

Fahrstreifen basiert, ähnlich wie Liang in [91]. Die Methode beruht auf folgenden Annahmen: das Fahrzeug fährt auf und parallel zur Fahrbahn, die horizontal ist und geradlinige parallele Fahrstreifenmarkierungen besitzt (bis zu einer Entfernung von ca. 50 m). Diese Annahmen sind meist in einem typischen Autobahn- oder Landstraßenszenario und oft in städtischen Szenarien ausreichend erfüllt.

Werden die Fahrstreifenmarkierungen im Bild erkannt, so können diese auf die Straßenebene anhand der bekannten Kameraparameter transformiert werden. Diese Transformation hängt von den Orientierungsparametern des Sensors beziehungsweise Sensorsystems ab, welche zu bestimmen sind. Dazu wird daher eine Optimierungssuche (Gauß-Newton'sche Methode) basierend auf dieser Transformation vorgeschlagen, die die Parallelität der Geraden auf der Straßenebene zueinander und zur Fahrtrichtung maximiert. Der Ablauf der Prozessschritte ist in Abbildung 4-6 graphisch dargestellt.

Diese Methode läuft komplett automatisch und benötigt einen sehr geringen Rechenaufwand. Die Methode ist in den meisten Fahrsituationen anwendbar und, im Unterschied zu anderen Verfahren (wie Liang [91]), kann ein Ausfall detektiert werden. Außerdem führt die vorliegende Methode eine Optimierung der extrahierten Beobachtungen durch. Ein weiterer Vorteil ist die Wiederverwendbarkeit der Module für eine Fahrstreifenenerkennung. Nachteilig sind die durch Abweichung von den Annahmen hinzugekommenen Ungenauigkeiten.

### 4.3.1 Extraktion der Fahrstreifenmarkierungen

Zur Extraktion der Fahrstreifenmarkierungen wird die in [25, 67] vorgestellte Methode des Autors verwendet. Die Methode ist unabhängig von der Bahn- oder Markierungsfarbe, der Markierungsdicke und der Fahrstreifenbreite. Obwohl sie für eine gerade Spur entwickelt wurde, werden auch Fahrbahnmarkierungen in leichten Kurven in der Nähe zum Fahrzeug erkannt.

In der Methode werden die mit der Hough-Transformation erhaltenen Bildgeraden  $[\theta_i \ \rho_i]^T$  (siehe Seite 36) betrachtet. Um Geraden wegzufiltern, die nicht zu den Fahrstreifenmarkierungen gehören (Rauschen oder andere Strukturen), werden sie auf die Straßenebene mittels der Matrix  $\mathbf{T}_r^i$  transformiert. Diese Transformation, die anhand der Alingment-Parameter durchgeführt wird, wird im Anhang B hergeleitet. Von den erhaltenen Geraden in Straßenebenekoordinaten  $[\theta_r \ \rho_r]^T$  werden nur die beibehalten, die parallel zur Fahrzeug- $X$ -Richtung verlaufen, also  $\theta_r = 0^\circ$ . Da Fahrzeugschwankungen stattfinden können, wird die Zulässigkeit der Filterung mit



$\theta_r \in (0 - p_{\theta_r}, 0 + p_{\theta_r})$  erhöht ( $p_{\theta_r}$  sei hier ein fixer Parameter). Damit sind beide Markierungen als kleines Bündel von  $k$  Geraden gegeben, die als Beobachtungen einer Zufallsgröße angesehen werden können – siehe Abbildung 4-7 oben.

### 4.3.2 Gauß-Newton'sche Suche der Orientierungsparameter

Die Ausrichtung der ermittelten Fahrstreifenmarkierungen kann als Funktion der Bildgeraden und der Orientierungsparameter ausgedrückt werden (ausführlich im Anhang B beschrieben):

$$\theta_r = f(\theta_i, \rho_i, t_x, t_y, t_z, \omega, \phi, \kappa) \quad (4.5)$$

Da sich die zu suchenden Fahrzeugschwankungen direkt auf die Kamera auswirken, können diese als (geringe) Abweichung der Kameraorientierung gesehen werden, so dass  $\omega = \omega_0 + d\omega$ ,  $\phi = \phi_0 + d\phi$  und  $\kappa = \kappa_0 + d\kappa$  gilt. Die innere Orientierung und die Alignment-Parameter der Kamera werden als konstant und bekannt angenommen.

Um die drei unbekannten Parameter  $d\omega$ ,  $d\phi$  und  $d\kappa$  zu bestimmen ( $p = 3$ ), wird eine Gauß-Newton'sche Suche durchgeführt. Dazu wird folgende zu minimierende Funktion verwendet, die auf die beobachteten  $k$  Bildgeraden angewendet wird:

$$\sum_{j=1}^k (\theta_{r,j} - \tilde{\theta}_{r,j})^2 \rightarrow \min! \quad (4.6)$$

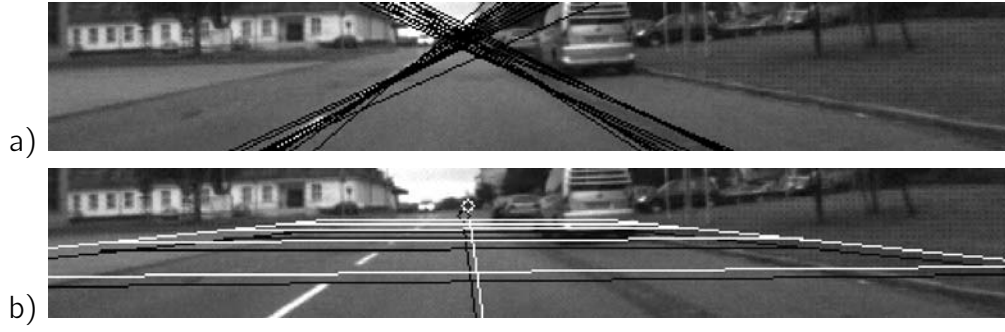
Diese entspricht der Summe der Fehlerquadrate der beobachteten Ausrichtung, wobei der erwartete Wert  $\tilde{\theta}_r$  für jede Beobachtung  $j$  als  $0^\circ$  vorgegeben ist (parallel zur Fahrtrichtung).

Damit das Minimierungsproblem lösbar ist, muss  $k > p$  sein. Ist  $k = p$ , so wird die exakte Lösung anhand der Beobachtungen berechnet (ohne Optimierung der Fehlerquadrate). Im Fall, dass eine Gerade pro Markierung extrahiert wird, können also zwei Parameter exakt daraus berechnet werden, und zwar  $d\phi$  und  $d\kappa$ , die den größten Einfluss auf die Beobachtungen haben. Der Rollwinkel  $d\omega$  wird dabei als konstant ( $d\omega = 0^\circ$ ) angenommen<sup>5</sup>. Sind mehrere Beobachtungen pro Markierung vorhanden, so können diese beiden Parameter im Sinne der Fehlerquadrate optimiert werden. Im Fall, dass lediglich Beobachtungen für eine Markierung vorhanden sind, kann der Algorithmus nicht verwendet werden.

Der zu bestimmende Parametervektor wird als  $\mathbf{a} = [d\phi \ d\kappa]^T$  ausgedrückt.

In jedem Iterationsschritt werden der Residuenvektor

<sup>5</sup>Um den Rollwinkel  $d\omega$  mitschätzen zu können, wären weitere Bedingungen in der Gauß-Newton'schen Suche notwendig, wie zum Beispiel gegeben durch senkrechte Elemente oder durch die Kenntnis der Fahrstreifenbreite.



**Abbildung 4-7:** Beispiel der Korrektur der Fahrzeugorientierung. a) Extrahierte Fahrstreifenmarkierungen. b) Vergleich der Ausrichtung mit Korrekturen (helles Gitter) und ohne (dunkles Gitter).

$$\mathbf{r} = \begin{bmatrix} \theta_{r,1} - \tilde{\theta}_{r,1} \\ \theta_{r,2} - \tilde{\theta}_{r,2} \\ \vdots \\ \theta_{r,k} - \tilde{\theta}_{r,k} \end{bmatrix} = \begin{bmatrix} \theta_{r,1} \\ \theta_{r,2} \\ \vdots \\ \theta_{r,k} \end{bmatrix} \quad (4.7)$$

und die Jacobi-Matrix

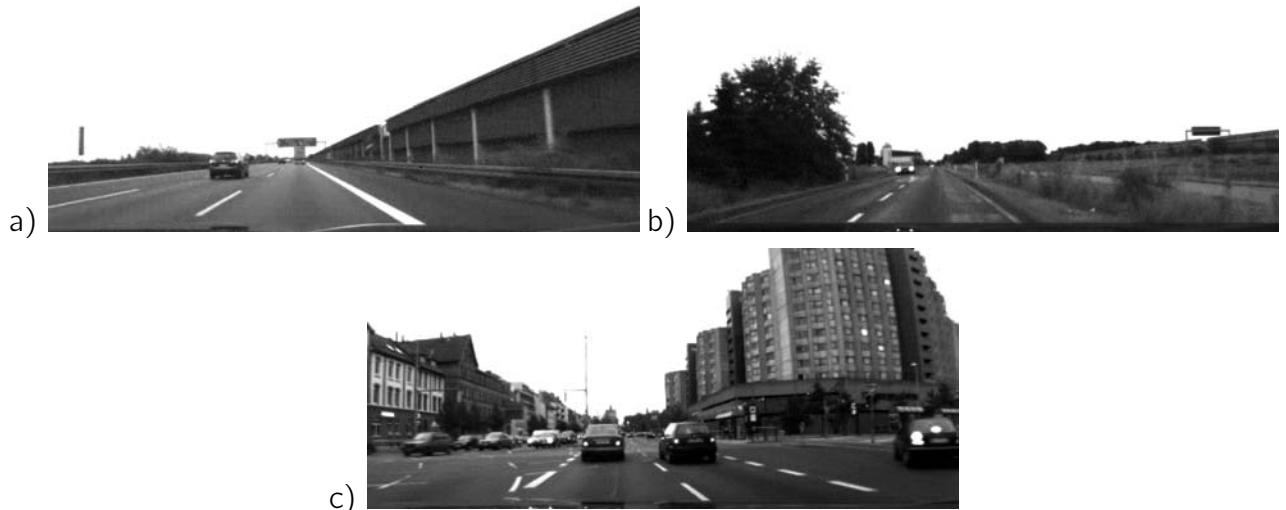
$$\mathbf{J} = \begin{bmatrix} \frac{\partial \theta_{r,1}}{\partial d\phi} & \frac{\partial \theta_{r,1}}{\partial d\kappa} \\ \frac{\partial \theta_{r,2}}{\partial d\phi} & \frac{\partial \theta_{r,2}}{\partial d\kappa} \\ \vdots & \vdots \\ \frac{\partial \theta_{r,k}}{\partial d\phi} & \frac{\partial \theta_{r,k}}{\partial d\kappa} \end{bmatrix} \quad (4.8)$$

aufgebaut. Die Berechnung dieser Matrix ist trivial aber sehr umfangreich und wird daher in diesem Text nicht weiter detailliert ausgeführt.

Ausgehend von einer guten Initialschätzung für die Parameter  $d\phi$  und  $d\kappa$  (in diesem Fall beide gleich  $0^\circ$ ) und von einem Endkriterium  $p_c$  wird folgender Iterationsschritt entsprechend der Gauß-Newton'schen Methode bis  $\|\mathbf{a}_{i+1} - \mathbf{a}_i\| < p_c$  fortgeführt:

$$\mathbf{a}_{i+1} = \mathbf{a}_i - (\mathbf{J}_i^T \cdot \mathbf{J}_i)^{-1} \cdot \mathbf{J}_i^T \cdot \mathbf{r}_i \quad (4.9)$$

Dieser Ausdruck entspricht der Normalgleichung in der Ausgleichungstheorie. Abbildung 4-7 zeigt die Korrekturen an einem Beispielbild. In Kapitel 9 wird auf die Validierung und die Tests des Verfahrens mit simulierten und realen Bildern eingegangen.



**Abbildung 4-8:** Beispiele von Fahrsituationen mit unterschiedlichem Fahrkorridor. a) Autobahn, b) Landstraße, c) Innenstadt.

## 4.4 Ermittlung des Fahrkorridors

Im Rahmen der Verarbeitung der Sensordaten auf Sensorebene soll nun ein Ansatz zur Bestimmung der Fahrkorridorbegrenzungen ermittelt werden. Diese kommen als Aufmerksamkeitssteuerung für die weitere Objekterkennung zum Einsatz. Die Ermittlung des Fahrkorridors wird durch sensornahe kooperative Fusion erreicht, in welche die Daten des Laserscanners, der Fahrstreifenerkennung (mit korrigierter Kameraausrichtung) sowie der Ortung und Informationen über die befahrene Straße aus einer digitalen Karte einfließen.

Der Fahrkorridor sei definiert als der lokale Bereich, in dem sich das Fahrzeug zum gegebenen Zeitpunkt befindet und einige weitere Zeitschritte danach befinden kann. Direkt benachbarte Verkehrsteilnehmer teilen denselben Fahrkorridor. Abbildung 4-8 stellt drei Beispielsituationen dar. In einem Autobahnszenario besteht der Fahrkorridor aus einem virtuellen entlang der Fahrtrichtung gezogenen Bereich, der alle Fahrstreifen dieser Fahrtrichtung umfasst. In einer Bundesstraße sind beide Fahrtrichtungen in dem Fahrkorridor enthalten. Im städtischen Bereich kann der Fahrkorridor neben der eigenen Fahrbahn Kreuzungen und Querstraßen beinhalten. Fußgängersteige (Gehkorridor) sind nicht Teil des Fahrkorridors.

Die Suche des Fahrkorridors ist in der Forschung ein aktuelles Thema zur Analyse der Fahrsituation. In der Literatur werden folgende Sensoren verwendet:

- **Mono-Kamera:** In den Kamerabildern sind meist die Straßenoberfläche, die Fahrstreifenmarkierungen bis hin zu seitlichen Elementen wie Leitplanken oder Vegetation sichtbar. Viele Arbeiten konzentrieren sich auf die Suche der Markierungen oder Begrenzungen der Fahrbahn. Beispielsweise stellen Loose und Franke [94] dazu einen Partikelfilter vor. Andere erkennen die Straßentextur im Bild, wie zum Beispiel Berger et al. [15]. Das Ziel dieser Systeme ist ein ungewolltes Verlassen des Fahrstreifens zu erkennen oder das autonome Fahren (Bahnplanung).
- **Ortung und digitale Karten:** In modernen, detaillierten, digitalen Straßenkarten sind Infor-

mationen vorhanden, die Aufschluss über den Fahrkorridor geben. Beispielsweise sind darin die Art von Straße (ob Autobahn, Bundesstraße oder Innenstadt), die Straßennetztopologie, die Anzahl der Fahrstreifen, deren Breite etc. enthalten.

- 3D-Sensor: In den Daten von 3D-Sensoren (wie Laserscanner, Radar oder Stereokameras) sind strukturelle Informationen über die Fahrsituation erhalten, denn es werden nicht nur Fahrzeuge und Hindernisse sondern auch die seitliche Bebauung oder Vegetation erfasst. Weiß et al. [158] verwenden ein *occupancy grid* anhand eines Laserscanners und Karteninformationen, um statische Objekte zu filtern und die Begrenzungen des Fahrkorridors in Abschnitte vor dem Fahrzeug zu schätzen. Darms et al. [38] verwenden in einem ähnlichen System ein scannendes Radar und eine Klothoidenschätzung. Loose und Franke [94] verwenden ebenso ein *occupancy grid* anhand einer Stereo-Kamera und eines Radars und verfolgen die Straßenränder anhand eines Partikelfilters und eines Kalman-Filters.

Alle diese Methoden liefern auf mehr oder weniger direkte Art Informationen über den Fahrkorridor oder dessen seitliche Begrenzungen. Die Zuverlässigkeit der einzelnen Sensoren ist jedoch in bestimmten Situationen eingeschränkt. Beispielsweise kann eine Kamera durch die Sonne geblendet werden oder die Fahrstreifenmarkierungen verschmutzt sein oder ganz fehlen. Ein Laserscanner kann schwer unterscheiden, ob ein langes seitliches Objekt durch einen mitfahrenden Lkw oder eine Leitplanke erzeugt wird. Durch kurzzeitige GPS-Aussetzer oder Ungenauigkeiten kann die Information über die befahrene Straße unsicher sein oder ausfallen. Digitale Karten können zudem veraltet sein.

Aus diesen Gründen bietet sich an, mehrere dieser Sensoren gemeinsam in einem Fusionssystem zu verwenden. Die hier für diesen Zweck entwickelte Methode basiert auf einer kooperativen Fusion, die den Suchbereich, in dem die Fahrkorridorbegrenzungen liegen können, minimiert. Diese Methode wurde zum Teil vom Autor in [26, 27] vorgestellt. Da der zu untersuchende Streckenabschnitt maximal bis 100 m vor dem Fahrzeug reicht, genügt es den Fahrkorridor als gerade und parallel zur Fahrtrichtung anzunehmen<sup>6</sup>. Damit ist er durch die linke und rechte Begrenzung beschrieben,  $c_l > 0$  und  $c_r < 0$  (*corridor*). Diese Angaben beziehen sich auf die Y-Achse in Fahrzeugkoordinaten<sup>7</sup>. Aus keinem der Sensoren können die genauen Begrenzungen des Korridors zuverlässig bestimmt werden. Durch den kooperativen Ansatz können diese den Suchbereich jedoch besser einschränken, in  $c_l \in [c_{l, \text{near}}, c_{l, \text{far}}]$  sowie  $c_r \in [c_{r, \text{far}}, c_{r, \text{near}}]$ . Dabei gibt es keinen Hauptsensor, womit Komplettausfälle unwahrscheinlicher sind. Zunächst werden die Begrenzungen initialisiert,  $c_{l, \text{near}} = 0$ ,  $c_{r, \text{near}} = 0$ ,  $c_{l, \text{far}} = +\infty$  und  $c_{r, \text{far}} = -\infty$ .

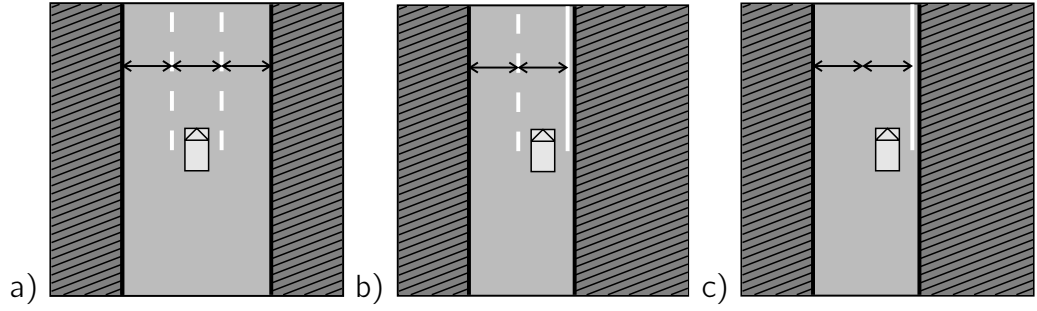
#### 4.4.1 Beitrag der Kameradaten zur Ermittlung des Fahrkorridors

Durch Kenntnis der relativen Position des Fahrzeugs zu den Fahrbahnmarkierungen und deren Art kann eine Aussage über die minimale Breite und Position des Fahrkorridors getroffen werden. Damit lassen sich die inneren Grenzen des Suchbereichs ( $c_{l, \text{near}}$  und  $c_{r, \text{near}}$ ) einschränken. Sind beide Markierungen gestrichelt, so wird angenommen, dass es mindestens einen weiteren (gleichbreiten) Fahrstreifen an jeder Seite geben muss. In diesem Fall beinhaltet der Korridor

---

<sup>6</sup>Mögliche Erweiterungen der Methode könnten eine andere Ausrichtung des Fahrkorridors, eine Krümmung oder auch Kreuzungsbereiche berücksichtigen.

<sup>7</sup>Nach links im positiven Sinne.



**Abbildung 4-9:** Beitrag der Kameradaten (Fahrstreifenmarkierungen) zur Ermittlung des Fahrkorridors. Im gestrichelten Suchbereich können die Begrenzungen des Fahrkorridors liegen.

mindestens drei Fahrstreifen. Ist eine der Markierungen durchgezogen, so wird dort die entsprechende Grenze des Suchbereichs angenommen.

$$c_{l,near} = \begin{cases} \max(c_{l,near}, l_l + w), & \text{wenn } l_l \text{ gestrichelt,} \\ \max(c_{l,near}, l_l), & \text{wenn } l_l \text{ durchgezogen.} \end{cases} \quad (4.10)$$

$$c_{l,far} = \max(c_{l,near}, c_{l,far}) \quad (4.11)$$

$$c_{r,near} = \begin{cases} \min(c_{r,near}, l_r - w), & \text{wenn } l_r \text{ gestrichelt,} \\ \min(c_{r,near}, l_r), & \text{wenn } l_r \text{ durchgezogen.} \end{cases} \quad (4.12)$$

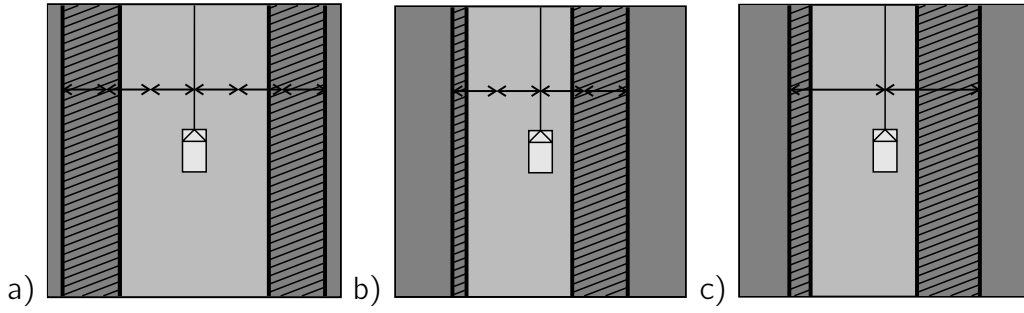
$$c_{r,far} = \min(c_{r,near}, c_{r,far}) \quad (4.13)$$

In den Gleichungen (4.10) bis (4.13) stellen  $l_l$  und  $l_r$  die Lage der Fahrstreifenmarkierungen (*lane*) dar.  $w$  ist die beobachtete Fahrstreifenbreite (*width*), die sich als  $l_l - l_r$  berechnen lässt. Abbildung 4-9 stellt drei Beispiele schematisch dar.

Die Fahrstreifenerkennung kann in manchen Situationen fehlen oder zu fehlerhaften Daten führen. Fehlt die Fahrstreifeninformation der Kamera, so wird von gestrichelten Markierungen und einer Standardbreite des Fahrstreifens von 3,5 m ausgegangen. Damit ist der Fahrkorridor wahrscheinlich zu breit geschätzt, was allerdings eine Sicherheit gegen das Verpassen von Objekten darstellt. Bei fehlerhaften Daten werden auf Basis der Fusion die nötigen Korrekturen stattfinden (siehe unten).

#### 4.4.2 Beitrag von Ortung und digitalen Karten zur Ermittlung des Fahrkorridors

Durch Kenntnis der befahrenen Straße und deren Eigenschaften lassen sich die äußeren Grenzen des Suchbereichs ( $c_{l,far}$  und  $c_{r,far}$ ) einschränken. Diese entsprechen der jeweils nach links und rechts vom Fahrzeug projizierten Straßenbreite:



**Abbildung 4-10:** Beitrag der Karteninformationen zur Ermittlung des Fahrkorridors. Im gestreiften Suchbereich können die Begrenzungen des Fahrkorridors liegen. a) Dreispurige Straße; b) Zweispurige Straße; c) Einspurige Straße.

$$c_{l, far} = \min(c_{l, far}, n_l \cdot w) \quad (4.14)$$

$$c_{l, near} = \min(c_{l, near}, c_{l, far}) \quad (4.15)$$

$$c_{r, far} = \max(c_{r, far}, -n_l \cdot w) \quad (4.16)$$

$$c_{r, near} = \max(c_{r, near}, c_{r, far}) \quad (4.17)$$

In den Gleichungen (4.14) und (4.16) stellen  $n_l$  die Anzahl der Fahrstreifen und  $w$  die Fahrstreifenbreite aus der digitalen Karte dar, siehe Abbildung 4-10.

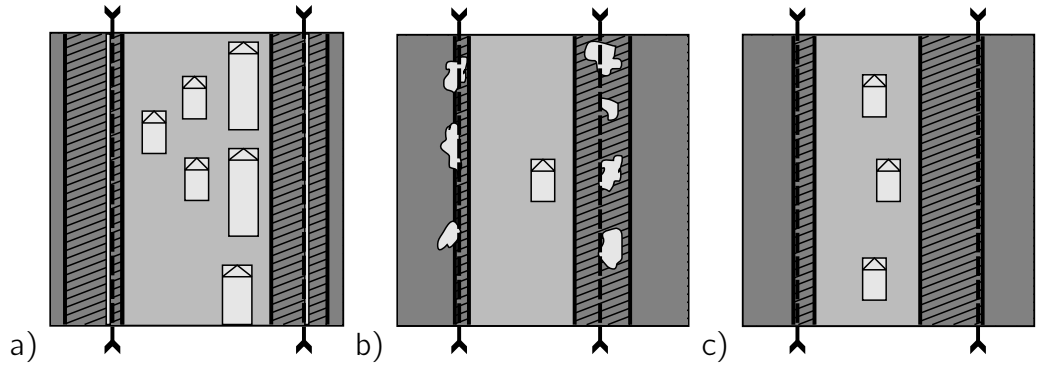
#### 4.4.3 Beitrag der Laserscannerdaten zur Ermittlung des Fahrkorridors

Die Laserscannerdaten beinhalten in vielen Fällen seitliche Straßenelemente. Diese stellen die Begrenzungen des Fahrkorridors dar. Um die Fahrkorridorbegrenzungen sicher zu bestimmen, müssen diese Elemente von anderen Verkehrsteilnehmern unterschieden werden. In diesem Ansatz wird dies mittels eines georeferenzierten, um das Fahrzeug lokalen, Gitters implementiert (*occupancy grid*), aus dem dynamische Objekte weggefiltert werden. In einem weiteren Schritt werden die möglichen Begrenzungen des Fahrkorridors daraus extrahiert.

Das georeferenzierte Gitter besteht aus einer kartesischen nordreferenzierten Matrix, in deren mittleren Position sich das Fahrzeug befindet. Die Größe einer Zelle bestimmt die Genauigkeit, mit der die Begrenzungen des Fahrkorridors erkannt werden können, aber auch den Rechenaufwand des Algorithmus. Da das Gitter lokal um das Fahrzeug liegt, muss es dessen Bewegung entsprechend in jedem Zeitschritt mitverschoben werden. Dies kann durch Anpassung der Matrixindizes als zweidimensionaler Ringpuffer sehr effizient implementiert werden.

Um das Gitter mit jedem Messzyklus zu aktualisieren, wird den Gitterpunkten entlang der Objektkonturen eine Menge  $p_+$  aufsummiert und anschließend allen Gitterpunkten eine Menge  $p_-$ , mit  $p_- < p_+$ , subtrahiert. Damit bleiben im Gitter nach mehreren Zyklen nur die Punkte ungleich Null, die zu statischen Objekten gehören.

Zur Extraktion der Fahrkorridorbegrenzungen aus dem Gitter wird eine vereinfachte Hough-



**Abbildung 4-11:** Beitrag der Laserscannerdaten zur Ermittlung des Fahrkorridors. Fahrkorridor wird gesetzt durch a) die Leitplanken (jedoch nicht durch die Lkws), b) Randbewuchs und c) die Obergrenze des Bereichs, mangels Laserscannerdaten.

Transformation verwendet (siehe Abschnitt 4.1 auf Seite 35). Der Vorteil der Hough-Transformation ist, dass selbst lückenhafte Strukturen gefunden werden können (wie im Falle einer Allee). Die gesuchten Geraden (Begrenzungen) bestehen aus einer Ausrichtung und einer Entfernung zum Ursprung im Gitter. Die Fahrtrichtung – korrigiert mit dem im letzten Unterabschnitt bestimmten Gierwinkel  $\kappa$  – wird bei der Suche vorgegeben. Damit muss nur der seitliche Abstand zum Fahrzeug (Ursprung) durch die Suche ermittelt werden,  $d_l$  und  $d_r$ .

Somit können die Begrenzungen des Fahrkorridors gesetzt werden, wie in Abbildung 4-11 beispielhaft gezeigt. Dabei werden die ermittelten Begrenzungen  $d_l$  und  $d_r$  nur angenommen, wenn diese im davor bestimmten Suchbereich liegen:

$$c_l = \begin{cases} c_{l,near}, & \text{wenn } d_l < c_{l,near}, \\ d_l, & \text{wenn } d_l \in (c_{l,near}, c_{l,far}), \\ c_{l,far}, & \text{wenn } d_l > c_{l,far}. \end{cases} \quad (4.18)$$

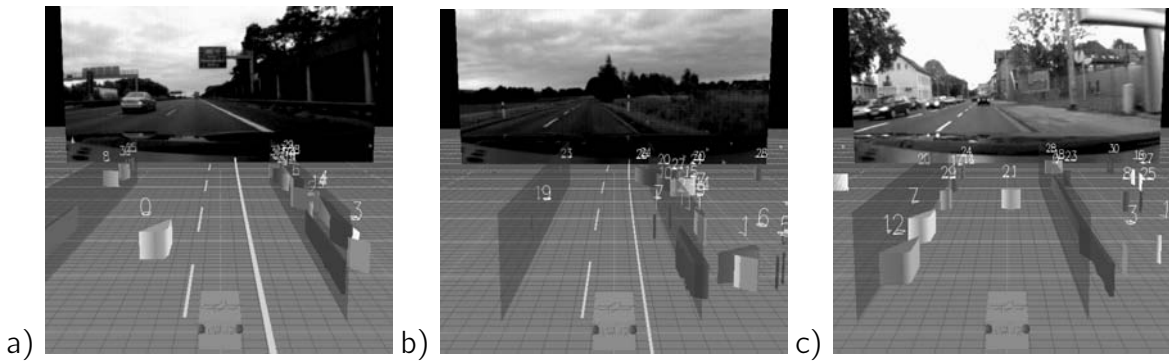
$$c_r = \begin{cases} c_{r,near}, & \text{wenn } d_r > c_{r,near}, \\ d_r, & \text{wenn } d_r \in (c_{r,far}, c_{r,near}), \\ c_{r,far}, & \text{wenn } d_r < c_{r,far}. \end{cases} \quad (4.19)$$

#### 4.4.4 Umgang mit widersprüchlichen Informationen

Mit dieser Methode werden wahrscheinliche Fahrkorridorbegrenzungen ermittelt. Durch die Einbindung von mehreren Sensoren ist die Wahrscheinlichkeit einer falschen Ausgabe gering.

Fällt einer der Sensoren aus, so wird durch die anderen eine Schätzung geliefert. Wenn beispielsweise keine Laserscannerdaten vorliegen, so werden die weitesten Grenzen des Suchbereichs übernommen, die aus den Fahrstreifendaten oder den digitalen Karten ermittelt wurden. Damit ist der notwendige minimale Erfassungsbereich, gegeben durch den Fahrkorridor, für die Objekterkennung sicher gestellt.

Sind jedoch die Informationen aus einem oder mehreren Sensoren falsch, so kann es zu Konflikten kommen. In solchen Fällen, wird immer die weiteste Grenze des Suchbereichs als Fahrkorridor-



**Abbildung 4-12:** Beispiele für die Ermittlung der Fahrkorridorbegrenzungen in unterschiedlichen Fahrsituationen, dargestellt mit dem Tool *AroundViewCar*.

begrenzung angenommen. Wenn beispielsweise beide Fahrstreifenmarkierungen gestrichelt sind, so muss der Fahrkorridor mindestens drei Fahrstreifen beinhalten, auch wenn die Karte nur zwei vorsieht. Auch beim Vorbeifahren an einem Lkw oder beim Aufenthalt neben anderen Objekten werden diese für statische Objekte gehalten und damit eine zu nahe Fahrkorridorbegrenzung durch den Laserscanner festgestellt. Diese Fälle werden durch die Einschränkungen aus den anderen Sensoren vermieden.

Zur Gesamteinschätzung dieses Verfahrens sind in Abbildung 4-12 einige Beispiele aufgeführt. In Kapitel 9 wird auf die Ergebnisse im Detail eingegangen.

## 4.5 Zusammenfassung des Kapitels

In diesem Kapitel wurde auf die Verarbeitung der Sensordaten für die Objekterkennung eingegangen. Für die Kamera wurden Methoden zur Kantenfilterung, Extraktion von Textur sowie Extraktion von Geraden vorgestellt.

Zusätzlich wurde auf die Problematik der Fahrzeugschwankungen während der Fahrt eingegangen. Zur Ermittlung dieser Fahrzeugschwankungen wurde eine neuartige Methode basierend auf den im Bild erkannten Fahrstreifenmarkierungen entwickelt. Diese werden automatisch erkannt, auf die Straßenebene transformiert und auf Parallelität mit der Fahrtrichtung geprüft. Da diese Transformation von den gesuchten Orientierungsparametern ( $\phi$  und  $\kappa$ ) abhängt, wird sie als Funktion für eine Gauß-Newton'sche Suche verwendet. Mit dieser Methode können die Sensordaten vor der eigentlichen Verarbeitung korrigiert werden.

Schließlich wurde auch ein Ansatz erörtert, mit dem der Fahrkorridor in der Fahrsituation durch sensornahe Fusion von Fahrstreifendaten, digitalen Karten und Laserscannerdaten robust erkannt werden kann. Das Neue an dieser Methode ist, dass sie als Fusion aus mehreren Sensordaten stattfindet, womit eine höhere Zuverlässigkeit erreicht wird. Außerdem wird hierbei ein einfaches Fahrkorridormodell verwendet, was eine schnelle Verarbeitung ermöglicht aber keine Begrenzungen in komplexen Situationen (wie enge Kurven und Kreuzungen) darstellen kann. Diese Methode wird für die Objekterkennung als Aufmerksamkeitssteuerung der Sensordaten verwendet, womit eine Reduktion der Datenmenge (Rauschobjekte am Straßenrand) und des Rechenaufwands erreicht werden kann.



## 5 Objekterkennung aus Bildern

In diesem Kapitel wird die Bildverarbeitung als Quelle zur Erzeugung von Objektbeobachtungen aus einzelnen Bildern vorgestellt. Hierzu werden die im letzten Kapitel vorverarbeiteten Kamerabilder und extrahierten Bildmerkmale verwendet. Die hier vorgestellten Methoden basieren auf dem Grauwerttexturbild und der bekannten Struktur der Bilder in einer Fahrsituation. Die hiermit erzeugten Beobachtungen werden in späteren Schritten verfolgt bzw. mit denen aus anderen Sensoren fusioniert (siehe Zusammenhang in Abbildung 5-1).

Zur Erzeugung von Objektbeobachtungen aus Bilddaten sind unterschiedliche Ansätze untersucht worden. Klassische Ansätze, wie sie im Kapitel 2 zum Stand der Technik vorgestellt werden, basieren auf angelernten Mustern, Konturdaten, Schattenzügen und der Bewegung im Bild. Alle diese Methoden weisen im Einzelnen wichtige Schwächen auf, die bei qualitativ beschränkten Bildern besonders kritisch sind.

Der hier entwickelte Ansatz, zum Teil vom Autor bereits in [28] vorgestellt, basiert auf der Segmentierung der Texturbilder und der Klassifikation der erhaltenen Bildsegmente nach deren Struktur. Der Vorteil gegenüber den oben erwähnten Ansätzen ist, dass die einzige Annahme zur Objekterkennung aus der Abhebung eines Objektes vom Hintergrund besteht. Diese fällt dem menschlichen Auge in Form von Licht und Schatten, Farbe und Textur oder Konturen und Linien auf. Die Verwendung eines geeigneten Texturindikators deckt jedoch alle diese Fälle ab. Die Herausforderung der Methode besteht in der Unterscheidung zwischen Objekten und anderen Bereichen oder Markierungen, was anhand eines regelbasierten Ansatzes gelöst wird.

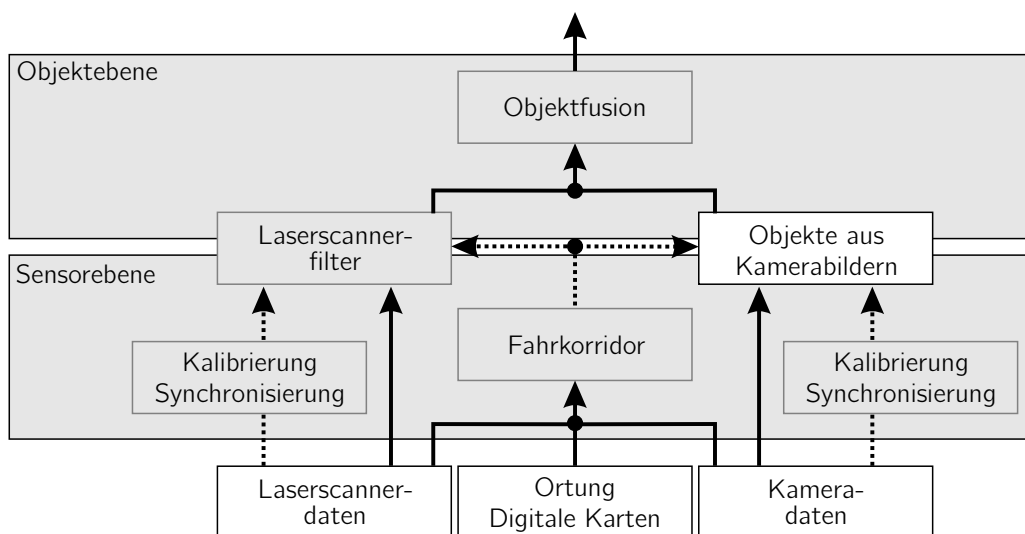
Zusätzlich werden die erkannten Objektbeobachtungen anhand von Kanteninformationen validiert und verfeinert, sowie mit einer angepassten Messunsicherheit, einer logischen Referenz und der partiellen Beobachtbarkeit versehen.

### 5.1 Extraktion der Bildstruktur

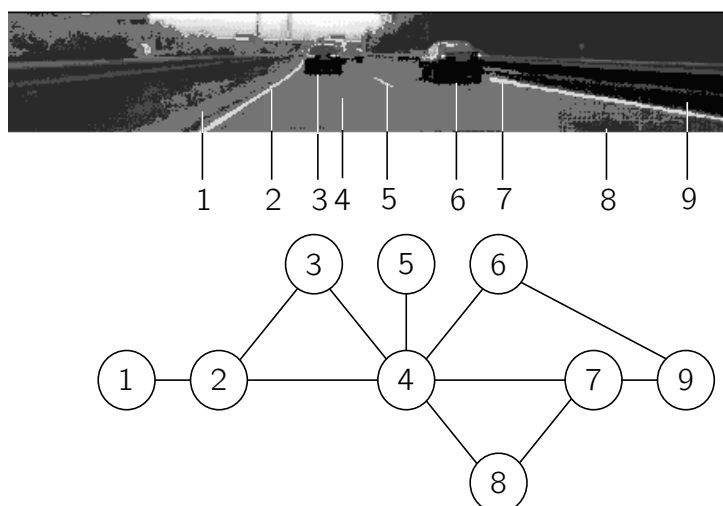
Die Struktur des im Kapitel 4.1 erhaltenen Texturbildes wird in einem Graphen berechnet, wie in Abbildung 5-2 exemplarisch dargestellt wird. Dabei gehören zusammenhängende Bereiche mit gleichem Wert zum selben Graphenknoten (Bildsegment). Dieser kann dem Bildhintergrund, der Straßenebene, einem Fahrzeug oder weiteren Objekten entsprechen. Jedes Segment besitzt einen Texturwert (Grauwert) und eine Liste von benachbarten Segmenten (Graphenkanten).

Beim Algorithmus zur Berechnung des Segmentgraphen wird das Bild durchlaufen und jedes Pixel wird mit den bereits betrachteten Nachbarpixeln verglichen (zum Beispiel das Linke und das Untere). Daraus können sich drei Situationen ergeben:

1. das Pixel hat den gleichen Grauwert wie eins der Nachbarpixel und gehört damit zum selben Bildsegment;
2. das Pixel hat einen anderen Grauwert als beide Nachbarpixel, womit ein neues Bildsegment initialisiert werden soll;



**Abbildung 5-1:** Objekterkennung aus Bildverarbeitung im Rahmen der Arbeit.



**Abbildung 5-2:** Ein vereinfachtes Texturbild und Ausschnitt des entsprechenden Bildsegmentgraphen.

3. das Pixel hat den gleichen Grauwert wie die beiden Nachbapixel, diese haben jedoch unterschiedliche Segmentnummern. Damit müssen die Bildsegmente vereinigt werden. Dieses bedeutet die Angleichung der Nachbarlisten und die Umsetzung aller Pixel des kleineren Bildsegments auf das andere, womit das kleinere Bildsegment gelöscht wird.

Die Bildränder werden in diesem Prozess als weitere Segmente (ohne Grauwert) behandelt (*bottom, top, left, right*).

Die Berechnung des Segmentgraphen ist eine sehr rechenintensive Aufgabe. Zur Beschleunigung wird nur innerhalb des Fahrkorridors (aus Kapitel 4.4) und unterhalb des Horizonts ausgewertet<sup>1</sup>. Außerdem wird die Reihenfolge des Algorithmus so gewählt, dass der mittlere Bereich (zumeist Straße) zuerst verarbeitet wird. Eine letzte Optimierung erfolgt durch die Verkleinerung des Eingangsbildes auf eine gröbere Auflösung. Im späteren Abschnitt 5.4 wird die Verfeinerung der erkannten Objektbeobachtungen mittels Kanteninformationen vorgestellt.

## 5.2 Klassifikation der Bildsegmente

Nachdem der Segmentgraph aufgestellt ist, werden die Bildsegmente klassifiziert. Jedes Bildsegment kann ein Teil der Straße, eine Markierung auf der Straße, ein Fleck bzw. Schatten auf der Straße, Teile eines Objekts auf der Straße, oder Randbereiche oder Objekte außerhalb der Straße sein.

$$Segment \in \{road, markers, spot, object, others\}$$

Die Segmentklassifikation erfolgt nun durch Anwendung eines Regelsatzes, der die typische Struktur eines Bildes in einer Fahrsituation darstellt. Die Regeln basieren auf folgenden grundsätzlichen Beziehungen zwischen Bildsegmenten:

1. Mögliche Klassen eines Bildsegments ( $PossibleClass(i,cl) \rightarrow \{true, false\}$ )
2. Benachbarte Bildsegmente ( $Neighbor(i,j) \rightarrow \{true, false\}$ )
3. Vergleich der Texturwerte von benachbarten Bildsegmenten ( $Lighter(i,j) \rightarrow \{true, false\}$ ,  $Equal(i,j) \rightarrow \{true, false\}$ ).

Initialisiert werden alle Bildsegmente als möglich zugehörig zu allen Klassen ( $PossibleClass(i,cl) = true \forall i,cl$ ). Das Referenzsegment, das zur Auslegung der Texturmerkmale herangezogen wurde (siehe Kapitel 4.1), wird als Straße gesetzt und dient für die Klassifikation des Graphen als Keim. Die Regeln sind in negierter Form ausgedrückt, so dass sie bestimmte Klassen für bestimmte Bildsegmente ausschließen. Bleibt für ein Bildsegment eine einzige mögliche Klasse übrig, so wird ihm diese zugeordnet. Der Regelsatz wird solange auf die nichtklassifizierten Bildsegmente angewendet, bis keine Regel mehr zutrifft. Tabelle 5-1 beinhaltet exemplarisch einen vereinfachten Regelsatz für die Analyse von Segmentgraphen.

<sup>1</sup>Die Transformation des Fahrkorridors in Bildkoordinaten erfolgt anhand der in Abschnitt 4.3 verwendeten und im Anhang B detailliert beschriebenen Transformation  $\mathbf{T}_r^i$ .

**Tabelle 5-1:** Beispiel eines Regelsatzes für die Klassifikation eines Bildsegments.

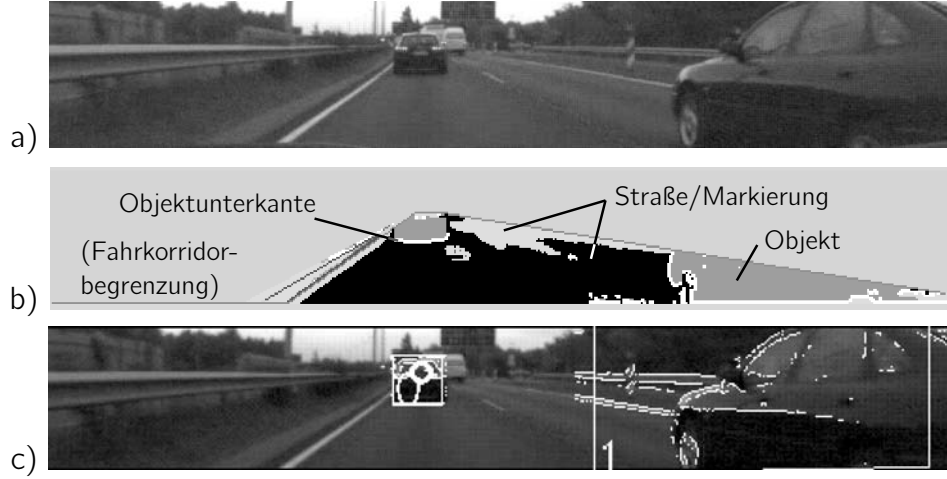
Regel	Bedingungen	Ausschluss
Ein Objekt (dessen schattiger unterer Bereich) ist in Kontakt mit einem helleren Straßensegment.	$\nexists j \mid Neighbor(i,j) \ \& \ PossibleClass(j,road) \ \& \ Lighter(j,i)$	<i>object</i>
Ein Segment in Kontakt mit der Bildunterkante gehört zur Straße, wenn ein Straßensegment in Kontakt mit der Bildunterkante den gleichen Grauwert besitzt.	$Neighbor(i,bottom) \ \& \ \nexists j \mid Neighbor(j,bottom) \ \& \ PossibleClass(j,road) \ \& \ Equal(i,j)$	<i>road</i>
Ein Segment, das durch eine Markierung von einem Straßensegment getrennt ist, muss denselben Wert haben, um zur Straße zu gehören.	$Neighbor(i,j) \ \& \ PossibleClass(j,marker) \ \& \ \exists k \mid Neighbor(j,k) \ \& \ PossibleClass(k,road) \ \& \ !Equal(i,k)$	<i>road</i>
Eine Straßenmarkierung ist in Kontakt mit einem dunkleren Straßensegment.	$\nexists j \mid Neighbor(i,j) \ \& \ PossibleClass(j,road) \ \& \ Lighter(i,j)$	<i>marker</i>
Ein Fleck auf der Straße ist in Kontakt mit nur einem helleren Straßensegment.	$\nexists j \mid Neighbor(i,j) \ \& \ PossibleClass(j,road) \ \& \ Lighter(j,i)$	<i>spot</i>

Wegen der hohen Komplexität der zu analysierenden Bilder kann jedem Bildsegment oft keine eindeutige Klasse zugeordnet werden. Solche Unsicherheiten entstehen beispielsweise bei Flecken auf der Straße, bei undeutlichem Übergang von Straße zu Objekt, oder bei unerwarteten Bildstrukturen, wie bei der Fahrt unter einer Brücke. Die unklaren Fälle werden an dieser Stelle nicht gelöst, sondern stattdessen als Objekte klassifiziert und durch die weiteren Verarbeitungsschritte gehandhabt.

Abbildung 5-3 zeigt ein Beispiel für ein segmentiertes Bild und die Klassifikation der Bildsegmente. Im segmentierten Bild ist der Bereich über dem Horizont sowie an den Korridorseiten ausgelassen. Die weißen Linien sind die Unterkanten der als Objekt klassifizierten Bildsegmente.

### 5.3 Erzeugung von Objektbeobachtungen

Aus als Objekt klassifizierten Bildsegmenten werden Objektbeobachtungen erzeugt. Für diesen Zweck wird das Segmentbild ein weiteres Mal durchlaufen. Die Pixel, die zu unteren Grenzen von Objektsegmenten (*object*) gehören, werden ausgewählt. Diese Punkte ergeben Linien (z.B. Konturen, Objektunterkanten oder Schattenzüge). Je nach Szenengeometrie kann jede dieser Linien der Unterkante eines oder mehrerer Objekte entsprechen, aber auch zum Teil Objektseiten oder andere Elemente beinhalten – siehe weiß markierte Linien in Abbildung 5-3.b. Aus diesem Grund wäre es nicht korrekt aus jeder solcher Linien eine Objektbeobachtung zu erzeugen. Stattdessen



**Abbildung 5-3:** Beispiel für die Klassifikation von Bildsegmenten: a) Eingangsbild; b) klassifizierte Segmente (Ausführung im Text); c) extrahierte Objektbeobachtungen.

werden nur Abschnitte, die im Bild näherungsweise horizontal erscheinen, ausgewählt und daraus eine Objektbeobachtung in Bildkoordinaten  $\mathbf{z}_{\text{cam},i}$  zwischen den Bildpunkten  $(x_{i,1}, y_{i,1})$  und  $(x_{i,2}, y_{i,2})$  generiert.

$$\mathbf{z}_{\text{cam},i} = [x_{i,1} \ y_{i,1} \ x_{i,2} \ y_{i,2}]^T \quad (5.1)$$

Damit entsteht eine Vielzahl an Objektbeobachtungen, die möglicherweise auch *false positives* beinhaltet. Auch die Verschmelzung und Zerlegung von Objekten (*split* und *merge*) sind nicht ausgeschlossen. Diese Effekte müssen in weiteren Schritten der Verarbeitung und Fusion behandelt werden.

Die Umrechnung eines Punktes von der Bildebene  $\mathbf{p}_i = [x_i \ y_i \ 1]^T$  in die Straßenebene<sup>2</sup>  $\mathbf{p}_r = [x_r \ y_r \ 1]^T$  erfolgt mit der Transformationsmatrix  $\mathbf{T}_i^r$  in homogenen Koordinaten, die im Anhang B detailliert beschrieben ist.

$$\mathbf{p}_r = \mathbf{T}_i^r \cdot \mathbf{p}_i \quad (5.2)$$

Die Transformation kann ebenso mit folgenden Funktionen ausgedrückt werden<sup>3</sup>:

$$x_r = f_{x_r}(x_i, y_i) = \frac{\mathbf{T}_i^r[1, 1] \cdot x_i + \mathbf{T}_i^r[1, 2] \cdot y_i + \mathbf{T}_i^r[1, 3]}{\mathbf{T}_i^r[3, 1] \cdot x_i + \mathbf{T}_i^r[3, 2] \cdot y_i + \mathbf{T}_i^r[3, 3]} \quad (5.3)$$

$$y_r = f_{y_r}(x_i, y_i) = \frac{\mathbf{T}_i^r[2, 1] \cdot x_i + \mathbf{T}_i^r[2, 2] \cdot y_i + \mathbf{T}_i^r[2, 3]}{\mathbf{T}_i^r[3, 1] \cdot x_i + \mathbf{T}_i^r[3, 2] \cdot y_i + \mathbf{T}_i^r[3, 3]} \quad (5.4)$$

<sup>2</sup>Die Straßenebene entspricht den Fahrzeugkoordinaten mit einer festen Höhe  $z = 0$ .

<sup>3</sup>Diese Ausdrücke sind ähnlich zu den klassischen Kollinearitätsgleichungen, beinhalten jedoch wichtige Anpassungen, die im Anhang B erläutert werden.

In dieser Transformation sind die korrigierten Kameraschwankungen für  $\phi$  (Nickwinkel) beinhaltet – siehe Abschnitt 4.3. Die Korrektur des Gierwinkels  $\kappa$  wird dabei nicht berücksichtigt. Ansonsten würden die Objekte in straßenparallelen Koordinaten erzeugt, die nicht immer mit den fahrzeugparallelen Koordinaten des Laserscanners übereinstimmen (z.B. beim Fahrstreifenwechsel).

Diese Transformation wird nun angewendet, um Objektbeobachtungen von Bildkoordinaten  $\mathbf{z}_{\text{cam},i}$  – siehe Gleichung (5.1) – in Fahrzeugkoordinaten  $\mathbf{z}_{\text{cam}}$  – siehe Kapitel 3 – zu transformieren. Dabei wird angenommen, dass  $\mathbf{z}_{\text{cam},i}$  der Objektunterkante entspricht und dass die Straße vor dem Ego-Fahrzeug eben ist, was in einer Entfernung von bis 50-100 m näherungsweise zutrifft. Für das Objekt in Fahrzeugkoordinaten wird als Referenzpunkt der Heckmittelpunkt  $(x_r, y_r)$  und die Breite ( $w$ ) genommen. Diese sind die einzigen beobachtbaren Komponenten, die einen Wert annehmen (die unbeobachtbaren sind im folgenden Ausdruck weggelassen worden):

$$\mathbf{z}_{\text{cam}} = \begin{bmatrix} x_r \\ y_r \\ w \end{bmatrix} = \begin{bmatrix} f_{x_r}\left(\frac{x_{i,1}+x_{i,2}}{2}, \frac{y_{i,1}+y_{i,2}}{2}\right) \\ f_{y_r}\left(\frac{x_{i,1}+x_{i,2}}{2}, \frac{y_{i,1}+y_{i,2}}{2}\right) \\ f_{y_r}(x_{i,1}, y_{i,1}) - f_{y_r}(x_{i,2}, y_{i,2}) \end{bmatrix} \quad (5.5)$$

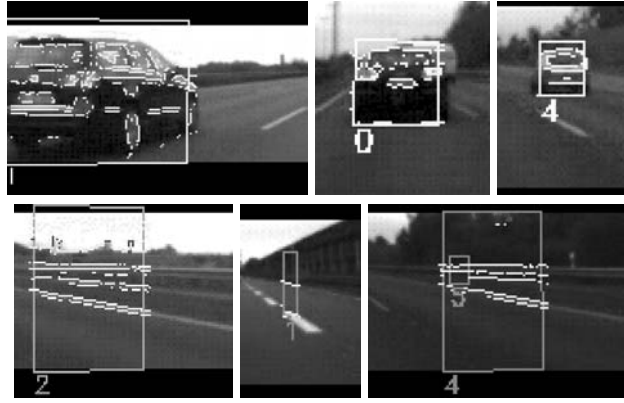
## 5.4 Validierung und Verfeinerung der Objektbeobachtungen

Da nicht alle generierten Objektbeobachtungen einem tatsächlichen Objekt entsprechen, werden zusätzliche Methoden verwendet, um diese automatisch zu validieren. Ähnliche Verfahren werden auch als sensor-nahe Fusion in der Literatur verwendet. Beispielsweise werden die Objektbeobachtungen aus einem Laserscanner oder Radar durch Bildverarbeitung validiert (bestätigt) bzw. verworfen (siehe Abschnitt 2.4). In diesem Fall werden die durch Bildverarbeitung gewonnenen Objektbeobachtungen anhand von anderen Bildverarbeitungsmethoden validiert und verfeinert.

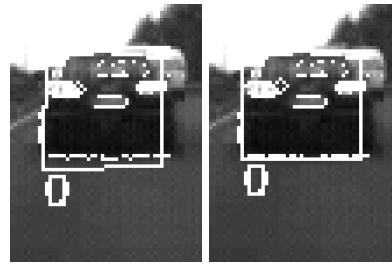
In dem gegebenen Bildausschnitt (Aufmerksamkeitssteuerung, *region of interest*) deuten viele Merkmale im Sinne der Validierung auf die Präsenz eines Objekts hin. Beispielsweise, haben Objekte eine andere Farbe als der Hintergrund, sie besitzen markante Punkte (Features) und Kanten, sie weisen Bewegung oder Symmetrie im Bild auf, etc. Aus diesen Kriterien zur Validierung der Objektbeobachtungen wird hier die Präsenz von Kanten im Bildausschnitt gewählt.

Vor der Validierung wird der zur Objektbeobachtung entsprechende Bildausschnitt berechnet. Die untere und die seitlichen Grenzen sind durch  $\mathbf{z}_{\text{cam},i}$  gegeben. Um die obere Grenze in Bildkoordinaten zu berechnen, wird eine Standardobjekthöhe  $h_v$  von 1,5 m für den Punkt  $(x_r, y_r)$  ins Bild berechnet:

$$\mathbf{p}_{i,h} = \mathbf{T}_v^i \cdot \begin{bmatrix} x_r \\ y_r \\ h_v \\ 1 \end{bmatrix} \quad (5.6)$$



**Abbildung 5-4:** Beispiele für die Validierung von Objektbeobachtungen. Die obere Reihe beinhaltet validierte und die untere nicht-validierte Objekte.



**Abbildung 5-5:** Beispiele für die Verfeinerung von Objektbeobachtungen aus Bildern. Links: Beobachtung aus dem Texturbild; Rechts: Verfeinerung mithilfe von Kantendaten.

Beinhaltet der Bildausschnitt wenige Kanten (Anzahl der Kantenpunkte durch Oberfläche ist kleiner als ein gegebener Parameter), so wird die Präsenz eines Objektes als unwahrscheinlich eingestuft. Die Kanten, die in Richtung des Fluchtpunktes liegen, werden davon ausgeschlossen, da sie auf kein Objekt hindeuten, sondern eher auf einen Bereich mit Straßenmarkierungen oder einen Seitenbereich (Leitplanken). Aus diesem Grund wird die Richtung der Kanten mit dem Fluchtpunkt  $\mathbf{p}_{vp}$  (*vanishing point*) verglichen. Siehe Abbildung 5-4 für validierte und nicht-validierte Beispiele.

Neben der Validierung einer Objektbeobachtung wird durch die Kanten auch die Position im Bild verfeinert. Da das Kantenbild feiner aufgelöst ist als das Texturbild, kann darin die Objektunterkante genauer erkannt werden. Für diesen Zweck wird um jeden Bildpunkt ein abstandsabhängiger Ausschnitt untersucht (nach Parameter  $p_{et}$ , *edge-texture*). Die Bildausschnitte lassen sich anhand dieses Parameters von der Straßenebene ins Bild berechnen:

$$\mathbf{T}_r^i \cdot \begin{bmatrix} x_r - p_{et} \\ y_r \\ 1 \end{bmatrix} \text{ und } \mathbf{T}_r^i \cdot \begin{bmatrix} x_r + p_{et} \\ y_r \\ 1 \end{bmatrix} \quad (5.7)$$

Für jedes Pixel zwischen  $(x_{i,1}, y_{i,1})$  und  $(x_{i,2}, y_{i,2})$  wird die Differenz  $c_{et}$  (*edge-texture correction*) zwischen der Kante  $y_e$  und der Texturgrenze  $y_t$  bestimmt und gemittelt,  $\overline{c_{et}}$  (sofern eine Kante an der Stelle vorhanden ist). Damit werden die Positionen  $y_{i,1}$  und  $y_{i,2}$  korrigiert und die Koordinaten auf der Straßenebene anhand der Gleichung (5.5) Neuberechnet. Abbildung 5-5 zeigt

das Ergebnis der Korrektur an einem Beispiel. Der Wert  $\overline{c_{et}}$  wird auch zur Einschätzung der Genauigkeit verwendet, wie im nächsten Abschnitt erläutert wird.

## 5.5 Genauigkeit und Beobachtbarkeit der Objektbeobachtungen

Die Genauigkeit der Objektbeobachtungen wird nun anhand der Varianzfortpflanzung von Bild- auf Straßenebene transformiert. Diese ermöglicht die Umrechnung der Kovarianzmatrix  $\sigma_x$  in einen neuen Raum  $\sigma_f$  anhand der Transformationsfunktion  $f$ . Für lineare Funktionen mit Design-Matrix  $\mathbf{A}$  gilt Folgendes:

$$\mathbf{x} \rightarrow f(\mathbf{x}) = \mathbf{A} \cdot \mathbf{x} \quad (5.8)$$

$$\sigma_x \rightarrow \sigma_f = \mathbf{A} \cdot \sigma_x \cdot \mathbf{A}^T \quad (5.9)$$

Ist die Funktion nichtlinear, so wird die Jacobi-Matrix mit den partiellen Ableitungen der Funktion aufgebaut, mit  $A_{ij} = \frac{\partial f_i(\mathbf{x})}{\partial x_j}$ .

Im Fall der Transformation zwischen Bild- und Straßenebene tritt eine nichtlineare Funktion auf (siehe Gleichungen (5.2) bis (5.4)), die von der Pixelposition abhängt. Zusätzlich zur Pixelposition wird an dieser Stelle der Einfluss einer Abweichung der Alignment-Parameter einbezogen. Insbesondere der Nickwinkel  $\phi$  und Gierwinkel  $\kappa$  variieren beim Fahren und haben einen Einfluss auf die Genauigkeit der Objektpositionen. Aus diesem Grund werden diese auch als Parameter der Transformation zur Varianzfortpflanzung berücksichtigt ( $\mathbf{T}_i^r(\phi, \kappa)$ ,  $f_{x_r}(x_i, y_i, \phi, \kappa)$  und  $f_{y_r}(x_i, y_i, \phi, \kappa)$  in Ausdruck (5.5)).

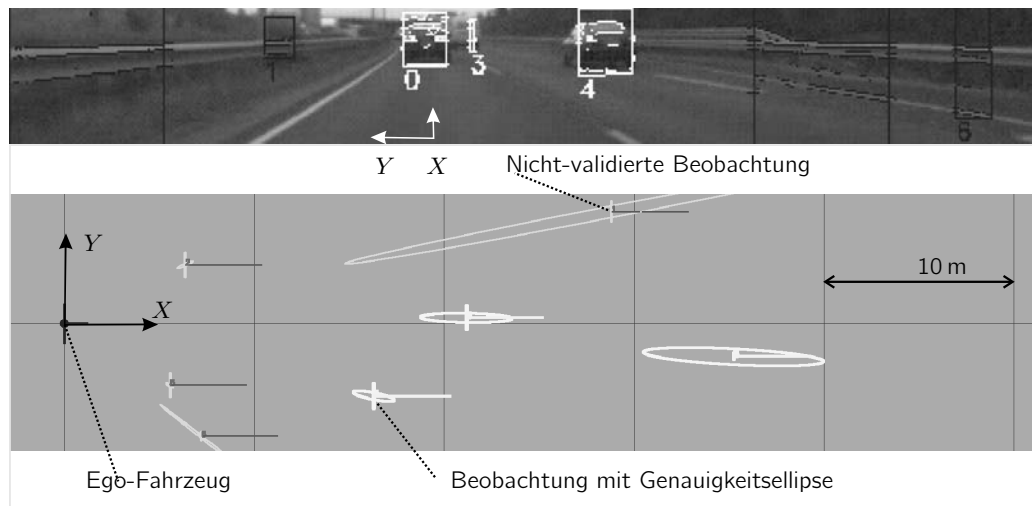
Die Berechnung der partiellen Ableitungen von  $\mathbf{T}_i^r$  kann einmalig pro Systemeinstellung bestimmt und festcodiert werden, so dass der Rechenaufwand gering bleibt.

Die Eingangskorrelationsmatrix in Bildkoordinaten entspricht folgender Diagonalmatrix:

$$\sigma_x = \text{Diag}(\sigma_{x_{i,1}}^2, \sigma_{y_{i,1}}^2, \sigma_{x_{i,2}}^2, \sigma_{y_{i,2}}^2, \sigma_\phi^2, \sigma_\kappa^2) \quad (5.10)$$

Dabei werden nur Varianzen und keine Korrelationen, wie z.B. zwischen horizontaler und vertikaler Bildkomponente, genommen. In den praktischen Arbeiten haben sich folgende Werte bewährt:  $\sigma_{x_i} = 4$  Pixel,  $\sigma_{y_i} = \overline{c_{et}}$  Pixel (Mittelwert der Korrekturen mit dem Kantenbild aus letztem Abschnitt; Wert liegt bei ca. 1,5 Pixel),  $\sigma_\phi = 0,5^\circ$  und  $\sigma_\kappa = 0,5^\circ$ . Die erhaltene Kovarianzmatrix wird als adaptive Messunsicherheit  $\mathbf{R}_{\text{cam}}$  in die Verfolgung bzw. Fusion eingehen. Abbildung 5-6 zeigt ein Beispiel mit erkannten Objekten im Bild und deren Transformation auf die Straßenebene. Die Kovarianz ist in Abhängigkeit der Position und der Korrektur mit Kanteninformation angepasst.





**Abbildung 5-6:** Räumlicher Vorwärtsschnitt und Genauigkeit der Objektbeobachtungen.

Für die weitere Verarbeitung bzw. Fusion der Daten aus der Bildverarbeitung wird angenommen, dass die Objekte lediglich von einer Seite erfasst werden<sup>4</sup> und dass deren Ausrichtung parallel zu der des eigenen Fahrzeugs ist<sup>5</sup>. Die logische Referenz für die weitere Verarbeitung entspricht damit dem Mittelpunkt der dem Sensor zugewandten Seite, hier bezeichnet als  $r = B$  (back). Die Länge und Ausrichtung des Objekts sind als nicht beobachtbar gesetzt. Da der Fokus dieser Arbeit auf der Position und Ausdehnung der Objekte auf der Straßenebene liegt, wird die Objekthöhe zunächst nicht betrachtet.

## 5.6 Zusammenfassung des Kapitels

Die vorgestellte Methode besteht in der Extraktion der Struktur des Texturbildes und der Anwendung von deskriptiven Regeln zur Klassifikation der Bildsegmente. Daraus lassen sich Bildsegmente hervorheben, die zu Objekten gehören.

Diese Methode ist für jeden Objekttyp gültig, da ausschließlich die Unterscheidung vom Hintergrund genutzt wird. Dank der adaptiven Texturberechnung ist die Methode von der Beleuchtung in der Fahrsituation unabhängig. Daher ist die Wahrscheinlichkeit, ein Objekt nicht zu detektieren (*false negative*), sehr gering.

Da die Erkennung sehr von der Struktur der Szene abhängt, hat diese Methode ihre Grenzen in Situationen, die sehr von der erwarteten Struktur abweichen. Beispielsweise bereiten die Durchfahrt unter eine Brücke (breite Schattenkante), die Präsenz von mehreren Texturen auf der Fahrbahn oder die Komplexität in einer innenstädtischen Situation noch Herausforderungen bezüglich der Leistung der Methode.

Eine quantitative Aussage über die Güte der Methode wird im Kapitel 9 anhand von experimentellen Ergebnissen ausgeführt.

<sup>4</sup>Diese Seite kann vorne, hinten, links und rechts vom tatsächlichen Objekt repräsentieren. Dass zum Teil zwei Seiten gleichzeitig erfasst werden können, wird hier vernachlässigt.

<sup>5</sup>Nach der Definition von Kapitel 3 kann die eigentliche Ausrichtung des Objekts um 90°, 180° oder 270° gedreht sein.



## 6 Objekterkennung aus Laserscannerdaten

In diesem Kapitel werden Verfahren zur Verarbeitung von Laserscannerdaten beschrieben, um Objektbeobachtungen zu extrahieren, die mit den Objektinformationen aus der Bildverarbeitung in die Objektfusion eingehen. Abbildung 6-1 zeigt den Zusammenhang dieses Kapitels in der Arbeit (vgl. Abbildung 1-3 auf Seite 7).

Ein Laserscanner liefert eine Liste von Rohpunkten oder Reflexionen pro Scan. Die Rohdaten müssen vorgefiltert, geometrisch referenziert, gruppiert und identifiziert werden.

Es wird zunächst die im Sensor stattfindende Segmentierung übernommen, die auf [59] basiert. Zur Wegfilterung von Rauschobjekten dienen Position- und Größenkriterien sowie der in Abschnitt 4.4 erkannte Fahrkorridor. Die Segmente werden dann an das in Abschnitt 3.2 eingeführte Objektmodell angepasst. Wie in Abschnitt 2.3.3 bereits erwähnt, besitzen die in der Literatur verwendeten Ansätze zur Modellanpassung, wie das Drei-Punkte-Modell, Einschränkungen bezüglich der Anpassung an runde oder konkave Objekte. Daher wird hier zur Modellanpassung eine neuartige Methode entworfen, die auf den inneren Winkel zwischen einem Konturpunkt und den Endpunkten basiert, um die Grundform eines Objekts zu entscheiden.

Zusätzlich werden in diesem Kapitel die optimale Objektreferenz, die partielle Beobachtbarkeit für die Laserscannerrohdaten sowie die angepasste Messunsicherheit der Objekte behandelt. Diese Methoden wurden zum Teil vom Autor in [28] vorgestellt.

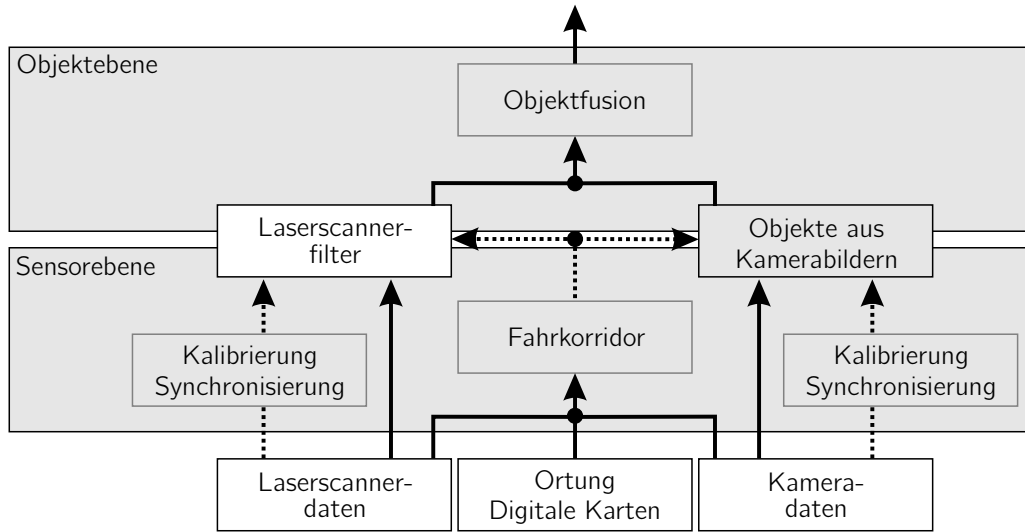
### 6.1 Anpassung an ausgerichtetes Quadermodell

Die Anpassung an das in Kapitel 3.2 beschriebene Quadermodell wird in drei Schritten durchgeführt. Zunächst wird die Grundform des Objekts bestimmt (gerader Teil, Eckteil, runder Teil oder konkaver Teil). Daraus werden dann die Ausdehnung und Ausrichtung des Objekts extrahiert, um die Struktur der Beobachtung  $\mathbf{z}_{ls} = [x \ y \ \dot{x} \ \dot{y} \ \varphi \ w \ l]^T$  zu füllen<sup>1</sup>. Schließlich wird die logische Referenz  $r$  des Objekts anhand seiner geometrischen Eigenschaften bestimmt.

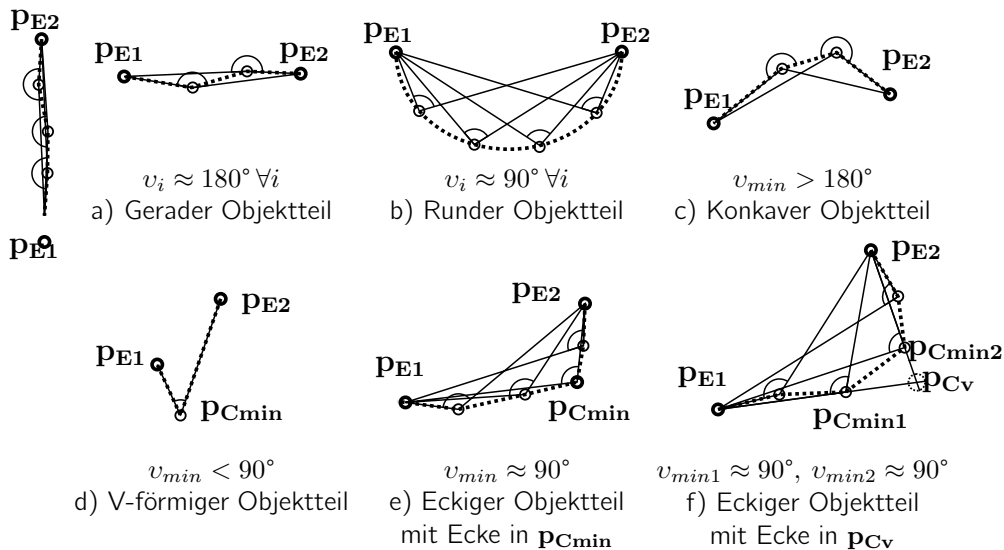
Zur Bestimmung der Grundform wird hier eine Methode entwickelt, die in der Lage ist, nicht nur gerade und eckige, sondern weiter differenziert auch runde und konkave Objekte zu charakterisieren. Die vorgestellte Methode betrachtet als allgemeiner Indikator für die Form eines Objektes den Winkel zwischen den beiden Endpunkten ( $\mathbf{p}_{E1}$  und  $\mathbf{p}_{E2}$ ) und jedem der inneren Konturpunkte ( $\mathbf{p}_{Ci}$ ), also  $v_i = \angle \mathbf{p}_{E1}, \mathbf{p}_{Ci}, \mathbf{p}_{E2}$  mit  $i \neq E1$  und  $i \neq E2$ . Insbesondere der Konturpunkt mit minimalem Winkel,  $v_{min} = \min(v_i)$ , ist aussagekräftig für den Konturtyp. Es wird dabei davon ausgegangen, dass die Objektpunkte in Erfassungsreihenfolge gegeben sind. Die möglichen Fälle, die sich in Abhängigkeit von  $v_{min}$  ergeben können, sind in Abbildung 6-2 dargestellt, wie gerader, runder, konkaver, v-förmiger und eckiger Objektteil. Für ein eckiges Objekt entspricht der Konturpunkt mit minimalen Winkel  $v_{min} \approx 90^\circ$  dem Eckpunkt. Ist das Minimum nicht deutlich, so wird der Eckpunkt wie in Abbildung 6-2.f verdeutlicht bestimmt.

---

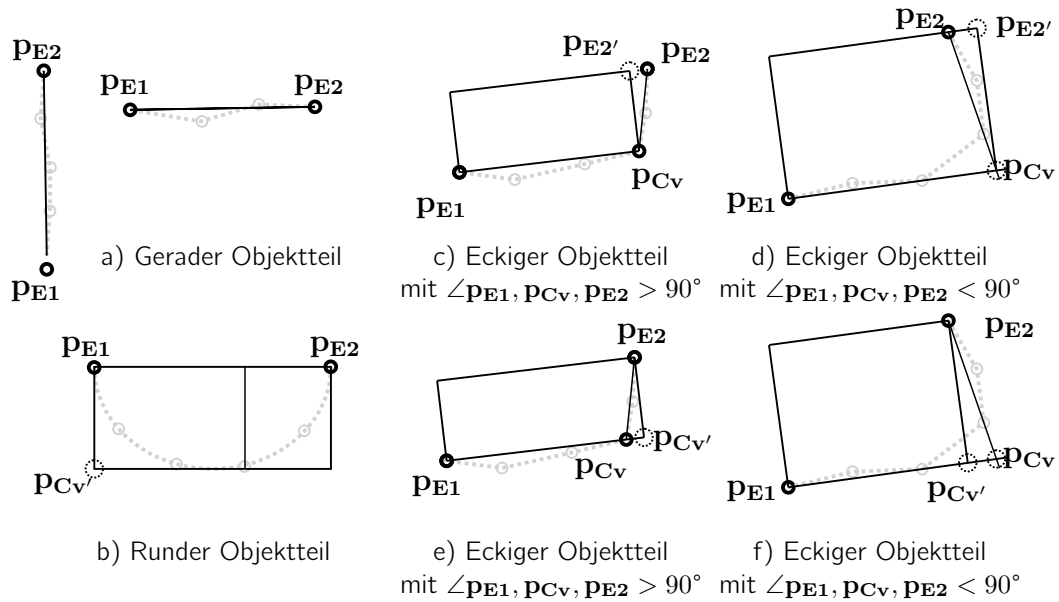
<sup>1</sup>Nur der beobachtbare Teil.



**Abbildung 6-1:** Objekterkennung anhand von Laserscannerdaten im Rahmen der Arbeit.



**Abbildung 6-2:** Grundform der Laserscannerobjekte je nach  $v_{min}$  auf orthographischer Sicht, angenommen der Sensor befindet sich auf der unteren Seite.



**Abbildung 6-3:** Anpassung von Objektteilen an gerichtete Quader.

Um die bestimmte Grundform an das gerichtete Quadermodell anzupassen, wird die Ausrichtung der längeren Seite  $\varphi_r \in (-180^\circ, +180^\circ]$  und die Dimensionen des Objekts anhand der Punkte  $p_{E1}$ ,  $p_{E2}$  und  $p_{Cv}$  berechnet. Dieser Schritt ist in Abbildung 6-3 schematisch dargestellt.

Für einen geraden oder runden Teil wird als erste Dimension  $d_1$  der Abstand zwischen beiden Endpunkten  $p_{E1}$  und  $p_{E2}$  verwendet. Die zweite Dimension ergibt sich als die maximale Entfernung zwischen einem der Konturpunkte und der Gerade, die über  $p_{E1}$  und  $p_{E2}$  verläuft. Ein konkaves Objekt wird am Eckpunkt in zwei Objekte geteilt, damit die Verwendung im Modell möglich wird.

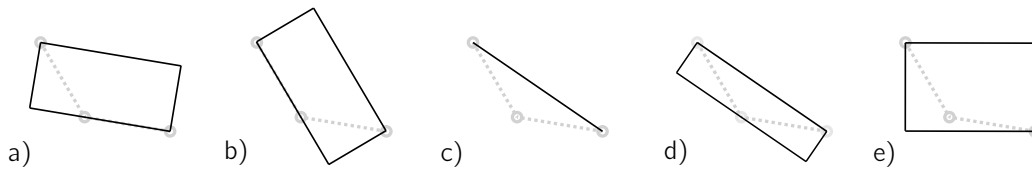
Im Fall eines eckigen oder v-förmigen Objektteils sind zwei Vorgehensweisen möglich. Bei der einen, in Abbildung 6-3.c und 6-3.d dargestellt, wird die längste Seite zwischen einem Endpunkt und dem ermittelten Eckpunkt (z.B.  $\overrightarrow{p_{Cv}p_{E2}}$ ) für die erste Dimension fixiert, während die zweite Dimension als Entfernung des anderen Endpunktes zu der ersten Seite berechnet wird. Wenn  $v_{min} > 90^\circ$  beträgt, wird die erste Dimension damit zu kurz geschätzt (6-3.c). Als zweite Möglichkeit, in Abbildung 6-3.e und 6-3.f dargestellt, wird der Eckpunkt auf der längsten Seite so verschoben, dass der resultierende Winkel mit beiden Endpunkten  $90^\circ$  beträgt. Ist der ursprüngliche Winkel kleiner als  $90^\circ$ , so wird die erste Dimension zu kurz geschätzt (6-3.f). Als optimale Lösung für einen eckigen Objektteil bietet sich die Verwendung beider Methoden in Abhängigkeit vom Winkel  $v_{min}$  an, und zwar die erste Methode wenn dieser unter  $90^\circ$  und die zweite wenn er über  $90^\circ$  beträgt.

Als letzter Schritt bei der Anpassung zum Modell  $z_{ls}$  wird die logische Referenz (definiert in Abschnitt 3.2 und Abbildung 3-1) des analysierten Objekts je nach seiner Ausrichtung, Lage, Form und Beobachtbarkeit gesetzt. Tabelle 6-1 stellt Beispiele für diese Regeln dar.

Die vorgeschlagene Vorgehensweise führt bei manchen Objektbeobachtungen zu Mehrdeutigkeiten, wie das Beispiel auf Abbildung 6-4 zeigt. An dieser Stelle können die Mehrdeutigkeiten ohne weitere Informationen nicht gelöst werden. Deshalb werden in solchen Fällen mehrere Objektbeobachtungen erzeugt und als alternativ markiert. In späteren Phasen der Verarbeitung, wenn die

**Tabelle 6-1:** Beispiele zur Entscheidung der Referenz eines Objekts anhand von deren Form, roher Ausrichtung  $\varphi_r$  und Position  $(x, y)$ . Das Ergebnis beinhaltet eine logische Referenz  $r$ , den entsprechenden Referenzpunkt  $\mathbf{p}_r$ , die resultierende Ausrichtung  $\varphi$  und die entsprechende Länge  $l$  und Breite  $w$  des Objekts.

Bedingungen	Form: Gerade $-45^\circ < \varphi_r \leq 45^\circ$ $y \geq 0$	Form: Ecke $-45^\circ < \varphi_r \leq 45^\circ$ $y < 0$	Form: Ecke $-135^\circ < \varphi_r \leq -45^\circ$ $x \geq 0$	Form: Rund/Konkav $-135^\circ < \varphi_r \leq -45^\circ$ $x \geq 0$
Ergebnis	$r=R$ $\mathbf{p}_r = \frac{(\mathbf{p}_{E1} + \mathbf{p}_{E2})}{2}$ $\varphi = \varphi_r$ $l = d_1; w = d_2$	$r=BL$ $\mathbf{p}_r = \mathbf{p}_{Cv}$ $\varphi = \varphi_r$ $l = d_1; w = d_2$	$r=BL$ $\mathbf{p}_r = \mathbf{p}_{Cv}$ $\varphi = \varphi_r + 90^\circ$ $l = d_2; w = d_1$	$r=BL$ $\mathbf{p}_r = \mathbf{p}_{Cv}$ $\varphi = \varphi_r + 90^\circ$ $l = d_2; w = d_1$
Darstellung				



**Abbildung 6-4:** Beispiel für Mehrdeutigkeiten in der Anpassung derselben Laserscanner-Reflexionen auf das gerichtete Quadermodell.

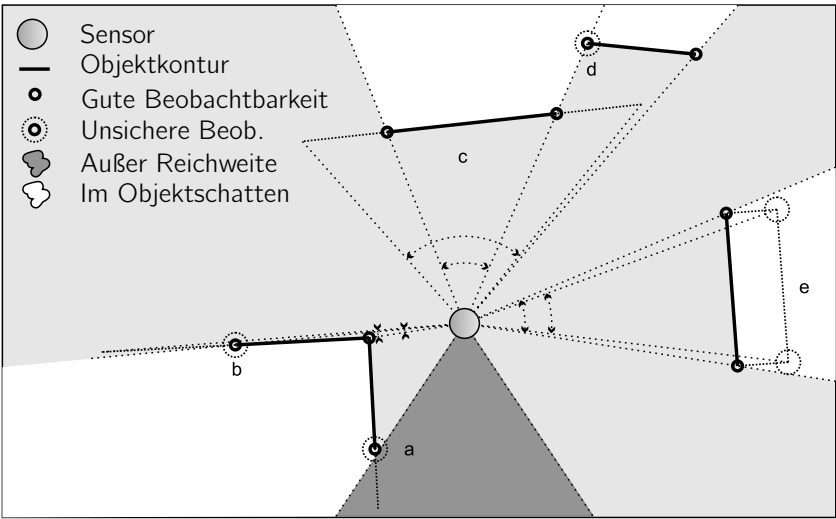
Bewegung der Objekte genauer bekannt ist, wird nur die passende Beobachtung berücksichtigt.

## 6.2 Beobachtbarkeit der Laserscannerobjekte

Da Objekte teilweise außerhalb des Sichtbereichs liegen oder sich gegenseitig verdecken können, wird hier deren Beobachtbarkeit geprüft. Diese Prüfung muss nur für die beiden Endkonturpunkte  $\mathbf{p}_{E1}$  und  $\mathbf{p}_{E2}$  durchgeführt werden. Anhand einfacher Plausibilitätskriterien wird überprüft, ob weitere Punkte hätten erfasst werden können und ob sie tatsächliche Endpunkte darstellen.

In folgenden Fällen wird ein Endpunkt eines Objektes als unsicher bzw. unbeobachtbar eingeschätzt:

1. Der Endpunkt liegt am Randbereich des Sichtfeldes des Laserscanners (Abbildung 6-5.a).
2. Der Endpunkt liegt im selben Winkel wie der Endpunkt eines anderen Objektes aber weiter entfernt vom Sensor, was auf eine mögliche Verdeckung hinweist (siehe Abbildung 6-5.d).
3. Unter Annahme, dass das Objekt doppelt so lang (bzw. breit) wäre, weist der Winkel aus dem Sensor zwischen erfasstem und fiktivem Endpunkt nur geringe Unterschiede auf. Damit ist die Winkelauflösung des Sensors zu grob für dieses Objekt (siehe Abbildung 6-5.b und 6-5.e als Beispiele und 6-5.c als Gegenbeispiel).



**Abbildung 6-5:** Bestimmung und Überprüfung der partiellen Beobachtbarkeit von Laserscanner-Objektbeobachtungen.

**Tabelle 6-2:** Beispiele für die Bestimmung der Beobachtbarkeit der Komponenten einer Objektbeobachtung und entsprechende Änderung der Referenz. Das Beibehalten der Referenz wird durch “=” symbolisiert.

$r$	$\mathbf{PE1}$ sicher?	$\mathbf{PE2}$ sicher?		$l$ beob.?	$w$ beob.?	neue $r$	neuer $\mathbf{p_r}$
$R$	ja	ja	$\Rightarrow$	ja	nein	=	=
$R$	ja	nein	$\Rightarrow$	nein	nein	$RB$	$\mathbf{PE1}$
$R$	nein	ja	$\Rightarrow$	nein	nein	$FR$	$\mathbf{PE2}$
$R$	nein	nein	$\Rightarrow$	nein	nein	=	=
$BL$	ja	ja	$\Rightarrow$	ja	ja	=	=
$BL$	ja	nein	$\Rightarrow$	ja	nein	=	=
$BL$	nein	ja	$\Rightarrow$	nein	ja	=	=
$BL$	nein	nein	$\Rightarrow$	nein	nein	=	=
...	...	...	...	...	...	...	...

Anhand der ermittelten logischen Referenz und der sicheren oder unsicheren Beobachtbarkeit der Endpunkte eines Objektes lässt sich auf die Beobachtbarkeit dessen Länge und Breite schließen. Aufgrund dessen wird der optimale Referenzpunkt beibehalten oder angepasst. Tabelle 6-2 zeigt einige Beispiele für diese Regeln, die sich in Verbindung mit den Beispielen in Tabelle 6-1 nachvollziehen lassen. Da die Geschwindigkeit der Objekte nicht direkt in den Laserscannerdaten erfasst wird, wird sie hier auch als unbeobachtbar gesetzt und erst bei der Objektverfolgung bestimmt.

### 6.3 Genauigkeit der Laserscannerbeobachtungen

Die bisherigen Betrachtungen in diesem Kapitel bezogen sich auf die im Sensor errechneten kartesischen Koordinaten der Reflexionspunkte. Die Rohdaten eines Laserscanners werden jedoch in Polarkoordinaten (mit Entfernung  $\rho$  und Winkel  $\theta$ ) erfasst. Jeder Messpunkt ist daher mit folgender Kovarianzmatrix charakterisiert:

$$\mathbf{\Sigma}_x = \begin{bmatrix} \sigma_\rho^2 & 0 \\ 0 & \sigma_\theta^2 \end{bmatrix} \quad (6.1)$$

Beide Komponenten werden als unkorreliert angenommen. Nach Erfahrungswerten beträgt  $\sigma_\rho = 0,5 \text{ m}$  und  $\sigma_\theta = 0,7^\circ$  im Hauptbereich und  $\sigma_\theta = 1,2^\circ$  im Seitenbereich (ab  $|\theta| > 30^\circ$ ).

Um diese Kovarianzmatrix in kartesische Koordinaten umzuwandeln, wird wie bei den Bildbeobachtungen die Varianzfortpflanzung verwendet – siehe Gleichungen (5.8) und (5.9) auf Seite 56. Die entsprechende Funktion für einen einzelnen Reflexionspunkt ist

$$\mathbf{f}(\rho, \theta) = \begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} \rho \cdot \cos(\theta + \theta_c) \\ \rho \cdot \sin(\theta + \theta_c) \end{bmatrix} \quad (6.2)$$

mit  $\theta_c$  als Nullausrichtung des Laserscanners.

Die Varianzfortpflanzung wird hier für den gesamten Beobachtungsvektor durchgeführt,  $\mathbf{z}_{ls} = [x \ y \ \varphi \ w \ l]^T$  (hier ohne Geschwindigkeitskomponenten). Dabei sind  $x$  und  $y$  die Koordinaten des Referenzpunktes  $\mathbf{p}_r$ . Die Komponenten  $w$  und  $l$  werden nur betrachtet, wenn sie beobachtbar sind – siehe letzten Abschnitt. Die Funktion für die Varianzfortpflanzung lautet somit:

$$\begin{bmatrix} x \\ y \\ \varphi \\ w \\ l \end{bmatrix} = \begin{bmatrix} f_x(\rho_r, \theta_r) \\ f_y(\rho_r, \theta_r) \\ \arctan\left(\frac{f_y(\rho_l, \theta_l) - f_y(\rho_v, \theta_v)}{f_x(\rho_l, \theta_l) - f_x(\rho_v, \theta_v)}\right) \\ \sqrt{(f_x(\rho_l, \theta_l) - f_x(\rho_v, \theta_v))^2 + (f_y(\rho_l, \theta_l) - f_y(\rho_v, \theta_v))^2} \\ \sqrt{(f_x(\rho_w, \theta_w) - f_x(\rho_v, \theta_v))^2 + (f_y(\rho_w, \theta_w) - f_y(\rho_v, \theta_v))^2} \end{bmatrix} \quad (6.3)$$



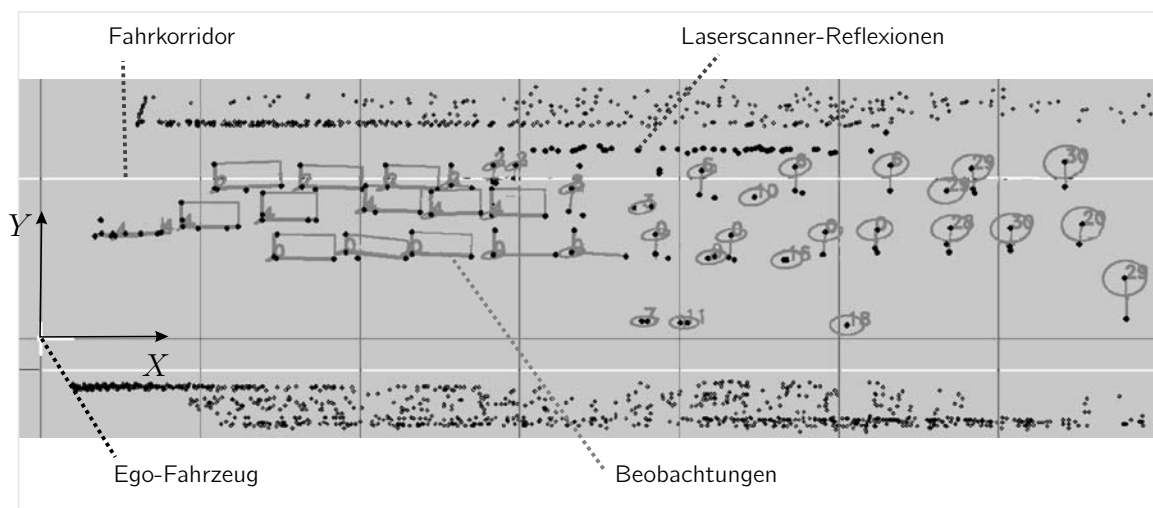
Dabei sind  $\rho_r$  und  $\theta_r$  die Rohkoordinaten des Referenzpunktes,  $\rho_v$  und  $\theta_v$  der Eckpunkt,  $\rho_l$  und  $\theta_l$  der Punkt, der die Länge und die Ausrichtung bestimmt (bspw.  $\mathbf{p}_{E1}$ ), und  $\rho_w$  und  $\theta_w$  der Punkt, der die Breite bestimmt (bspw.  $\mathbf{p}_{E2}$ ). Diese Punkte können je nach Objekttyp und Beobachtbarkeit dieselben sein. Für die Anwendung der Varianzfortpflanzung auf Ausdruck (6.3) wird die Design-Matrix  $\mathbf{A}$  und deren partiellen Ableitungen (Jacobi-Matrix) berechnet. Die damit erhaltene Kovarianzmatrix wird später in die Verfolgung bzw. Fusion als adaptive Messunsicherheit  $\mathbf{R}_{Is}$  eingehen.

## 6.4 Zusammenfassung des Kapitels

Der in diesem Abschnitt vorgestellte Prozess dient der Erzeugung von Objektbeobachtungen anhand der Einzelmesspunkte eines Scans. Nach der im Sensor stattfindenden Segmentierung und der Filterung anhand des Fahrkorridors werden die Objekte an ein gerichtetes Quadermodell angepasst. Dafür wird eine neuartige Methode anhand der inneren Winkel der Konturpunkte vorgestellt, um Objektformen differenziert klassifizieren (gerade, eckig, rund, konkav...) und charakterisieren zu können. Die Objekte werden zudem mit einer Referenz versehen, so dass sie trotz ihrer Ausdehnung durch einen Punkt (und somit durch das Modell aus Kapitel 3) dargestellt werden können. Auch die Beobachtbarkeit jeder Komponente eines Objektes wird hier anhand von unterschiedlichen Kriterien untersucht. Die Behandlung dieser Fragestellungen stellt einen neuartigen Aspekt gegenüber anderen Arbeiten aus diesem Themenbereich dar.

Die Vorteile dieses Verfahrens sind, dass es sehr detaillierte und trotzdem kompakte Informationen über die Objekte ergibt, und dies bei geringem Rechenaufwand. Ein Nachteil, den die meisten Ansätze aus der Literatur auch aufweisen, ist, dass Mehrdeutigkeiten bezüglich der Objektausrichtung entstehen können. Diese werden hier allerdings mitberücksichtigt und in späteren Schritten der Verarbeitung gelöst.

Damit sind die vorgestellten Methoden eine neuartige Zusammenstellung, um differenzierte Informationen über die Objekte zu bekommen. Die aus Laserscanner und Bildverarbeitung abgeleiteten Objektbeobachtungen sind vergleichbar und können miteinander fusioniert werden. Auf eine qualitative und quantitative Aussage über die Güte des Verfahrens wird im Kapitel 9 eingegangen. Abbildung 6-6 stellt ein Beispiel in einer Sequenz von Realdaten dar.



**Abbildung 6-6:** Beispiel der Erzeugung von Objektbeobachtungen anhand von Laserscannerdaten in einer Sequenz.

## 7 Objektverfolgung und Fusion

Dieses Kapitel befasst sich mit der Verfolgung und Fusion der aus den Sensordaten erzeugten Objektbeobachtungen. Es handelt sich dabei um eine konkurrierende Fusion auf Objektebene mit eingehenden Objektbeobachtungen gleicher Art. Zum Kern der Fusion wird der Informationsfilter verwendet. Der Fokus liegt dabei einerseits auf der Anpassung der Filterstrukturen an das verwendete Objektmodell (eingeführt in Abschnitt 3.2), einschließlich der variablen Referenz und partiellen Beobachtbarkeit. Andererseits wird auf Strategien zur Behandlung von typischen Herausforderungen der Objektverfolgung eingegangen, wie z.B. manövrierenden Objekten, Ausreißern sowie *split*- und *merge*-Effekten. Diese Strategien wurden vom Autor zum Teil in [29] veröffentlicht.

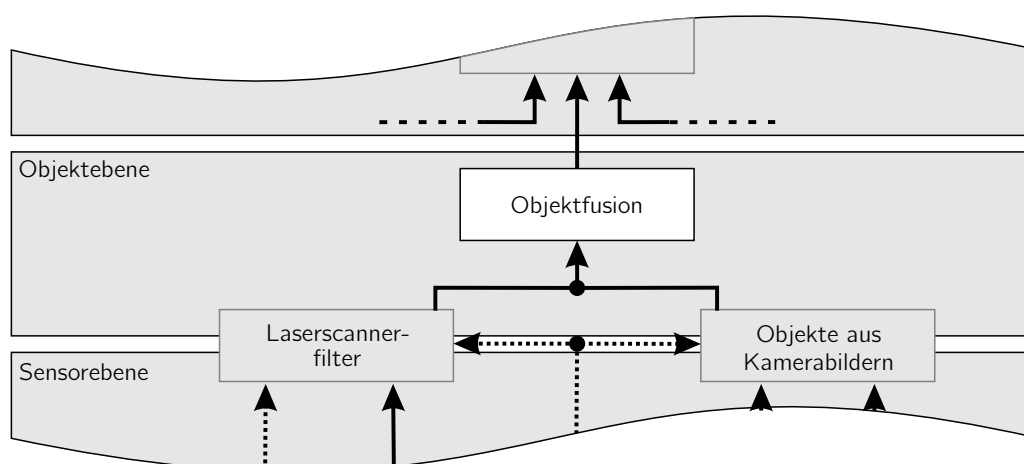
Abbildung 7-1 zeigt den Zusammenhang dieses Kapitels mit den Objekterkennungsmethoden und der restlichen Arbeit.

Der verwendete Ansatz stützt sich auf das im Stand der Technik vorgestellte klassische Multi-Target-Tracking-Schema (Abschnitt 2.2). Abbildung 7-2 stellt das angepasste Prozessmodell mit den entsprechenden Kapitelabschnitten dar. Dabei können die Beobachtungen von einem oder mehreren Sensoren (wie Kamera und Laserscanner) stammen.

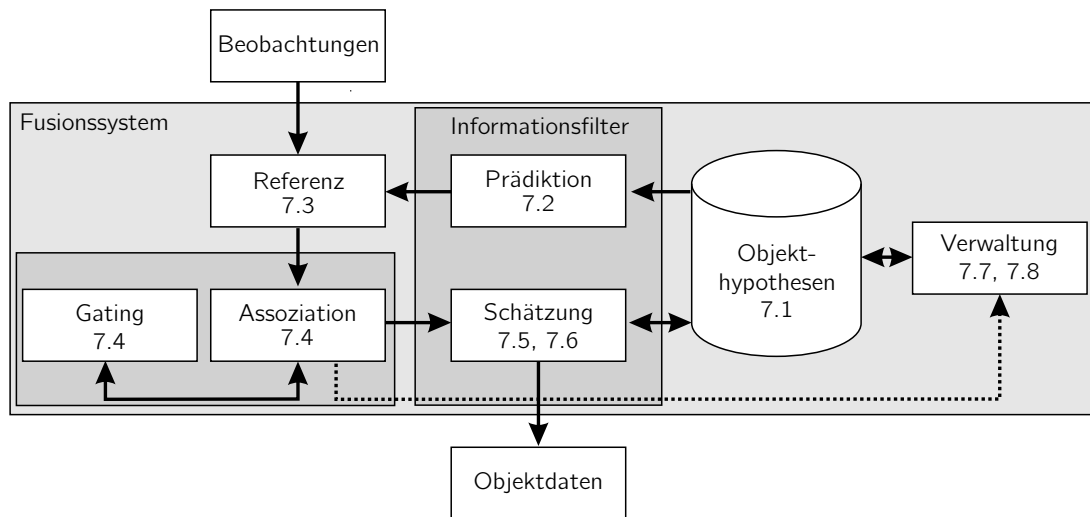
Zunächst wird in Abschnitt 7.1 auf das Dynamikmodell für die Objekte und die Herleitung der Systemmatrix sowie des Prozessrauschens eingegangen. Es wird hier ein einfaches geradliniges Modell mit konstanter Geschwindigkeit ausgewählt, das für die Filterstabilität vorteilhaft ist. Um Effekten entgegenzuwirken, die durch dieses Modell nicht abgedeckt sind, werden weitere Strategien entwickelt, wie im Folgenden beschrieben.

In jedem neuen Zyklus wird die Prädiktion jeder im System vorhandenen Objekthypothese anhand des Dynamikmodells und der Beobachtungsmatrizen durchgeführt. Der Zeitpunkt ist dank einer hardwarenahen Lösung hochgenau bekannt. Dieser Schritt wird in Abschnitt 7.2 beschrieben.

In Abschnitt 7.3 werden Mechanismen zum Wechseln der logischen Referenz eines Objektes als



**Abbildung 7-1:** Objektverfolgung und Fusion im Rahmen der Arbeit.



**Abbildung 7-2:** Prozessmodell für das Verfolgungs- und Fusionsmodul zur Übersicht des Kapitels.

Vorschrift zur Assoziation eingeführt. Damit werden die optimale Referenz und die entsprechenden Punkte der Objekthypothese und der Beobachtung berechnet.

Nachdem die prädizierten Objekthypothesen und die Objektbeobachtungen auf denselben Koordinatenraum und dieselbe logische Referenz gebracht sind, werden diese zueinander assoziiert (Abschnitt 7.4). Wichtig in diesem Schritt ist, dass Ausreißer und unklare Assoziationen berücksichtigt werden. Dazu werden Abstandsmaße und eine Zuordnungsstrategie definiert. Zusätzlich werden durch Gating-Mechanismen aus physikalischer oder statistischer Sicht unplausible Zuordnungen ausgeschlossen.

Anhand der Assoziationen werden im Schätzungsschritt die Objekthypothesen erneut geschätzt bzw. aktualisiert. Dafür wird in dieser Arbeit ein asynchroner Informationsfilter pro Objekthypothese verwendet, der sich für die konkurrierende Fusion der Daten aus Sensoren mit unterschiedlichen und variierenden Takten eignet. Der Informationsfilter wird in Abschnitt 7.5 detailliert vorgestellt. Des Weiteren werden in Abschnitt 7.6 dessen Ausdrücke an die Beobachtbarkeitsmatrix angepasst, so dass nur die beobachtbaren Komponenten (und deren Fehlerkovarianzen) geschätzt werden.

In den letzten Abschnitten 7.7 und 7.8 werden die Zuverlässigkeit der Objekthypothesen eingeschätzt und Verwaltungsaufgaben der Objekthypothesen durchgeführt. Darin müssen neue Objekte initialisiert und obsoletere Objekte aus der Liste entfernt werden. Außerdem werden auch Strategien vorgestellt, mit denen den negativen Effekten von manövrierenden Objekten, von *split*- und *merge*-Effekten und von Ausreißern entgegengewirkt wird. Dazu zählen die mehrfache Assoziation und die Verdoppelung sowie Vereinigung von Objekthypothesen.

Da es sich bei diesen Prozessschritten um einen geschlossenen Kreis handelt, findet nach dem letzten Schritt wieder die Prädiktion der Objekthypothesen statt.

## 7.1 Dynamikmodell

Die Zustandsgleichung (ohne Steuereingang) eines linearen dynamischen Systems lautet (siehe Anhang C.1)

$$\mathbf{x}(t_k) = \mathbf{F}(t_k) \cdot \mathbf{x}(t_{k-1}) + \mathbf{w}(t_k) \quad (7.1)$$

mit dem Zustandsvektor  $\mathbf{x}$ , der Übergangsmatrix  $\mathbf{F}$  und dem Prozessrauschen  $\mathbf{w}$ . Das hier verwendete Objektmodell für die verfolgten Objekthypothesen (eingeführt und begründet in Abschnitt 3.2) entspricht einer geradlinigen Bewegung mit konstanter Geschwindigkeit:

$$\mathbf{x}(t_k) = [x \ y \ \dot{x} \ \dot{y} \ \varphi \ w \ l]^T \quad (7.2)$$

Entsprechend Kapitel 3 gehören dazu die Fehlerkovarianzmatrix  $\mathbf{P}(t_k)$ , die logische Referenz  $r(t_k)$  und die Beobachtbarkeitsmatrix  $\mathbf{V}_x(t_k)$ . Die Übergangsmatrix  $\mathbf{F}(t_k)$  wird für einen beliebigen Zeitschritt  $\Delta t$  zu:

$$\mathbf{F}(t_k) = \begin{bmatrix} 1 & 0 & \Delta t & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \Delta t & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (7.3)$$

Als weißes Prozessrauschen werden hier indirekte Störeffekte aus im Modell nicht berücksichtigten Objektbeschleunigungen ( $\ddot{x}$ ,  $\ddot{y}$ ), der Drehrate ( $\dot{\varphi}$ , als Veränderung des Kurswinkels) sowie Veränderungen der Ausdehnung<sup>1</sup> ( $\dot{w}$ ,  $\dot{l}$ ) betrachtet. Der entsprechende indirekte Störvektor  $\mathbf{w}'$  (siehe Gleichung (7.4)) wird als weißes Rauschen mit der Kovarianzmatrix  $\mathbf{Q}'$  modelliert. Dessen Zusammenhang mit dem hier benötigten Prozessrauschvektor ist durch die Gewinnmatrix  $\mathbf{\Gamma}$  gegeben als:

$$\mathbf{w}(t_k) = \begin{bmatrix} w_x \\ w_y \\ w_{\dot{x}} \\ w_{\dot{y}} \\ w_{\dot{\varphi}} \\ w_w \\ w_l \end{bmatrix} = \mathbf{\Gamma}(t_k) \cdot \mathbf{w}'(t_k) = \begin{bmatrix} \frac{1}{2}\Delta t^2 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2}\Delta t^2 & 0 & 0 & 0 \\ \Delta t & 0 & 0 & 0 & 0 \\ 0 & \Delta t & 0 & 0 & 0 \\ 0 & 0 & \Delta t & 0 & 0 \\ 0 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 0 & \Delta t \end{bmatrix} \cdot \begin{bmatrix} w_{\ddot{x}} \\ w_{\ddot{y}} \\ w_{\dot{\varphi}} \\ w_{\dot{w}} \\ w_{\dot{l}} \end{bmatrix}, \quad (7.4)$$

<sup>1</sup>Durch unterschiedliche Objektkonstellationen können sich die erfassten Komponenten der Ausdehnung ändern.

Die Prozesskovarianzmatrix entspricht damit (Varianzfortpflanzung, vgl. [10])

$$\begin{aligned} \mathbf{Q}(t_k) &= \mathbf{\Gamma}(t_k) \cdot \mathbf{Q}'(t_k) \cdot \mathbf{\Gamma}^T(t_k) \\ &= \begin{bmatrix} \frac{1}{4}\Delta t^4 \sigma_{\ddot{x}}^2 & 0 & \frac{1}{2}\Delta t^3 \sigma_{\ddot{x}}^2 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4}\Delta t^4 \sigma_{\ddot{y}}^2 & 0 & \frac{1}{2}\Delta t^3 \sigma_{\ddot{y}}^2 & 0 & 0 & 0 \\ \frac{1}{2}\Delta t^3 \sigma_{\ddot{x}}^2 & 0 & \Delta t^2 \sigma_{\ddot{x}}^2 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2}\Delta t^3 \sigma_{\ddot{y}}^2 & 0 & \Delta t^2 \sigma_{\ddot{y}}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \Delta t^2 \sigma_{\dot{\varphi}}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \Delta t^2 \sigma_{\dot{w}}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \Delta t^2 \sigma_i^2 \end{bmatrix} \quad (7.5) \end{aligned}$$

Dabei wird  $\mathbf{Q}'$  als Diagonalmatrix  $\text{Diag}(\sigma_{\ddot{x}}^2, \sigma_{\ddot{y}}^2, \sigma_{\dot{\varphi}}^2, \sigma_{\dot{w}}^2, \sigma_i^2)$  modelliert (keine Korrelationen zwischen den Störeffekten). In den praktischen Versuchen im Rahmen dieser Arbeit haben sich die Werte  $\sigma_{\ddot{x}} = 0,005 \text{ m/s}^2$ ,  $\sigma_{\ddot{y}} = 0,005 \text{ m/s}^2$ ,  $\sigma_{\dot{\varphi}} = 0,002 \text{ }^\circ/\text{s}$ ,  $\sigma_{\dot{w}} = 0,001 \text{ m}$  und  $\sigma_i = 0,001 \text{ m}$  als sinnvoll erwiesen.

Mit dem modellierten Prozessrauschen lassen sich Fehler des vereinfachten Dynamikmodells (z.B. Beschleunigung, Kurven etc.) sowie der vereinfachenden Modellannahmen (z.B. Fahrzeuge sind keine Quader) zum Teil abfangen. Sind die Abweichungen von den Annahmen zu groß, so treten Assoziationsprobleme und weitere negative Effekte in der Verfolgung auf. Daher müssen weitere Fusionstrategien verwendet werden, wie im weiteren Verlauf des Kapitels vorgestellt.

Da in dieser Arbeit auf der Objektebene fusioniert wird, besitzen die eingehenden Beobachtungen dasselbe Format wie die verfolgten Objekthypothesen. Der Messvektor (jeweils  $\mathbf{z}_{\text{cam}}$  und  $\mathbf{z}_{\text{ls}}$ ) hat dieselben Komponenten wie der Zustandsvektor  $\mathbf{x}$ . Dazu gehören die jeweiligen Kovarianzmatrizen bzw. Messunsicherheiten ( $\mathbf{R}_{\text{cam}}$  und  $\mathbf{R}_{\text{ls}}$ ), die logischen Referenzen ( $r_{\text{cam}}$  und  $r_{\text{ls}}$ ) und deren Beobachtbarkeitsmatrizen ( $\mathbf{V}_{\mathbf{z}_{\text{cam}}}$  und  $\mathbf{V}_{\mathbf{z}_{\text{ls}}}$ ), die in den Kapiteln 5 und 6 im Detail erläutert wurden. Die Beobachtungsmatrizen ( $\mathbf{H}_{\text{cam}}$  und  $\mathbf{H}_{\text{ls}}$ ) entsprechen im Filter der Einheitsmatrix.

## 7.2 Prädiktion

Bei der Ankunft neuer Sensordaten besteht der erste Schritt in jedem Prozesszyklus in der Prädiktion der im System vorhandenen Objekthypothesen zu dem Zeitpunkt der Beobachtungen und der Abbildung in den Beobachtungsraum.

Zuerst werden alle Objekthypothesen anhand des soeben vorgestellten Dynamikmodells prädi-  
ziert,

$$\hat{\mathbf{x}}(t_k|t_{k-1}) = \mathbf{F}(t_k) \cdot \hat{\mathbf{x}}(t_{k-1}|t_{k-1}) \quad (7.6)$$

$$\mathbf{P}(t_k|t_{k-1}) = \mathbf{F}(t_k) \cdot \mathbf{P}(t_{k-1}|t_{k-1}) \cdot \mathbf{F}^T(t_k) + \mathbf{Q}(t_k) \quad (7.7)$$

mit Übergangsmatrix  $\mathbf{F}$  aus Gleichung (7.3) und Prozessrauschen  $\mathbf{Q}$  aus Gleichung (7.5).

Des Weiteren werden die prädizierten Objekthypothesen und deren Kovarianzmatrizen in den Beobachtungsraum transformiert, so dass die Assoziation dort stattfinden kann. Dank der Fusion auf Objektebene sind der Zustands- und der Beobachtungsraum äquivalent, womit diese Transformation trivial<sup>2</sup> ist. Der Vollständigkeit halber werden hier die Ausdrücke gegeben:

$$\hat{\mathbf{z}}(t_k|t_{k-1}) = \hat{\mathbf{z}}_s(t_k|t_{k-1}) = \mathbf{H}_s(t_k) \cdot \hat{\mathbf{x}}(t_k|t_{k-1}) \quad (7.8)$$

$$\mathbf{P}_{\hat{\mathbf{z}}}(t_k|t_{k-1}) = \mathbf{P}_{\hat{\mathbf{z}}_s}(t_k|t_{k-1}) = \mathbf{H}_s(t_k) \cdot \mathbf{P}(t_k|t_{k-1}) \cdot \mathbf{H}_s^T(t_k) \quad (7.9)$$

## 7.3 Variable Referenz und Referenzpunkt

In dieser Arbeit wird jedes Objekt (Objekthypothese und Objektbeobachtung) von einer logischen Referenz  $r$  begleitet, die besagt, auf welchen Punkt des Objekts sich der Objektzustand bezieht. Die Verwendung der variablen Referenz hat den Vorteil, dass das Modell punktbasiert und ohne Redundanzen bleibt, während gleichzeitig die dreidimensionalen Informationen berücksichtigt werden können.

Um Objekte vergleichen zu können, wird als Vorschritt zur Assoziation das Alignment durchgeführt. Hierbei wird für jede Kombination zwischen einer prädizierten Objekthypothese und einer Objektbeobachtung die optimale gemeinsame Referenz bestimmt sowie die entsprechenden Referenzpunkte und die Fehlerkovarianzmatrizen berechnet (siehe Abbildung 7-2). Dabei bleiben die Objekte unverändert. Soll die Kombination von Objekthypothese und Beobachtung in der Assoziation zugeordnet werden, so werden die bestimmte Referenz, der berechnete Referenzpunkt und die Kovarianzmatrix für die neue Schätzung verwendet.

### 7.3.1 Wechsel der Referenz eines Objekts

In Abhängigkeit von den beobachtbaren Komponenten eines Objekts (Länge und Breite) lässt sich seine Referenz wechseln (siehe Abbildung 7-3 links). Die Berechnung des Objektpunktes für eine bestimmte Referenz entspricht einer einfachen geometrischen Transformation, nämlich einer Translation um  $dx_{r \rightarrow r'} \in \{+l, +l/2, -l/2, -l\}$  und  $dy_{r \rightarrow r'} \in \{+w, +w/2, -w/2, -w\}$  (je nach ursprünglicher Referenz  $r$  und gezielter Referenz  $r'$ ) sowie einer zentrierten Rotation von

<sup>2</sup>Die Beobachtungsmatrix entspricht der Einheitsmatrix.

$\varphi$  um die  $Z$ -Achse:

$$x_{r'} = x_r + dx_{r \rightarrow r'} \cdot (\cos \varphi) - dy_{r \rightarrow r'} \cdot (\sin \varphi) \quad (7.10)$$

$$y_{r'} = y_r + dx_{r \rightarrow r'} \cdot (\sin \varphi) + dy_{r \rightarrow r'} \cdot (\cos \varphi) \quad (7.11)$$

Da sich die Fehlerkovarianzmatrix auf den Referenzpunkt bezieht, müssen die Komponenten für  $x$  und  $y$  anhand der Varianzfortpflanzung (siehe Abschnitt 5.5) ebenso angepasst werden. Dabei wird die Transformation nicht über die Länge bzw. Breite und die Ausrichtung durchgeführt (wie durch Peilen), weil dadurch zusätzliche Unsicherheiten bei jedem neuen Wechsel hinzugefügt würden<sup>3</sup>. Stattdessen wird die Transformation zur neuen Position anhand der Polarkoordinaten durchgeführt (siehe Ausdrücke aus Abschnitt 6.3),

$$\mathbf{f}(x, y) = \begin{bmatrix} f_x(f_\rho(x, y) + \delta\rho, f_\theta(x, y) + \delta\theta) \\ f_y(f_\rho(x, y) + \delta\rho, f_\theta(x, y) + \delta\theta) \end{bmatrix} \quad (7.12)$$

mit  $\delta\rho$  und  $\delta\theta$  berechnet aus  $x_r$ ,  $y_r$ ,  $x_{r'}$  und  $y_{r'}$  aus Gleichungen (7.10) und (7.11):

$$\delta\rho = f_\rho(x_{r'}, y_{r'}) - f_\rho(x_r, y_r) \quad (7.13)$$

$$\delta\theta = f_\theta(x_{r'}, y_{r'}) - f_\theta(x_r, y_r) \quad (7.14)$$

Daraus lässt sich als Design-Matrix  $\mathbf{A}$  die Jacobi-Matrix  $\mathbf{J}$  mit den partiellen Ableitungen von (7.12) am Betrachtungspunkt  $x_r$ ,  $y_r$  aufstellen und die neuen Positionskovarianzen (für Komponenten  $x$  und  $y$ ) bestimmen. Die restlichen Komponenten von  $\mathbf{P}$  bleiben unverändert (Siehe schematisches Beispiel in Abbildung 7-3 rechts).

### 7.3.2 Optimale gemeinsame Referenz von zwei Objekten

Der Algorithmus zur Bestimmung der optimalen gemeinsamen Referenz zwischen zwei Objekten besteht aus zwei Schritten. Zuerst werden jeweils die möglichen Referenzen herausgefunden. Dieses hängt von der Referenz der Objekte und von der Beobachtbarkeit<sup>4</sup> jeder Objektseite ab. Abbildung 7-3 links verdeutlicht diesen Schritt anhand eines Beispiels.

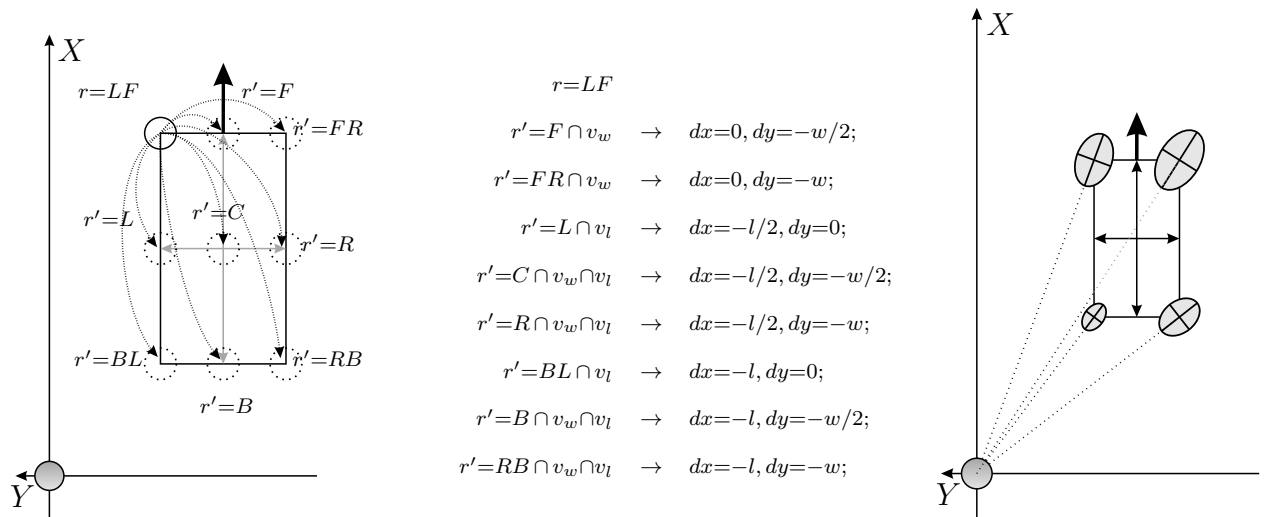
Im zweiten Algorithmusschritt wird die Referenz ausgewählt, die für beide Objekte möglich ist, und für die die Assoziationsdistanz (siehe nächsten Abschnitt) zwischen den entsprechenden Referenzpunkten minimal wird<sup>5</sup>. Ist keine gemeinsame Referenz möglich, so wird  $C$  (*center*) ausgewählt. Dieser unwahrscheinliche Fall entspricht, dass das Objekt von völlig unterschiedli-

<sup>3</sup>Selbst wenn zur alten Referenz gewechselt wird, erhöht sich die Ungenauigkeit.

<sup>4</sup>Im Abschnitt 7.6 ausführlich betrachtet.

<sup>5</sup>Möglicherweise ist der Abstand für mehrere Referenzen kaum unterschiedlich. In solchen Fällen wird die am nächsten zum Ego-Fahrzeug liegende Referenz genommen, da sie wahrscheinlich am besten beobachtet wird.





**Abbildung 7-3:** Links: Beispiel zum möglichen Wechseln der Referenz eines Objektes in Abhängigkeit der Beobachtbarkeit seiner Komponenten ( $v_w$ ,  $v_l$ ). Die Kreise stellen mögliche Referenzen des Objekts dar. Rechts: Veränderung der Objektgenauigkeit nach Referenz. Die Ellipsen stellen den Genauigkeitsbereich dar.

chen Seiten erfasst wird.

Ein Beispiel zur variierenden Referenz wird in Abbildung 7-6 auf Seite 85 zusammen mit der Beobachtbarkeit vorgestellt. Das entfernte Objekt wird zuerst als Punkt erfasst, womit die Referenz  $C$  beträgt. Zum Zeitpunkt  $t_1$  ist seine Breite bereits beobachtbar, womit  $BL$ ,  $B$  und  $RB$  mögliche Referenzen für Objekt und Beobachtung sind. Da alle drei praktisch denselben Abstand betragen, wird  $BL$  gewählt, denn der entsprechende Referenzpunkt liegt am nächsten zum Ego-Fahrzeug. Zum Zeitpunkt  $t_2$  wird das Objekt von der Seite erfasst, womit die einzige gemeinsame Referenz  $BL$  ist, die auch gewählt wird. Im letzten Zeitpunkt  $t_3$  wird nur noch die vordere Ecke beobachtet. Da sowohl die Länge als auch die Breite des Objekts beobachtbar sind, kann als Referenz  $LF$  gewählt werden.

## 7.4 Assoziation und Gating

Assoziation ist die Zuordnung zwischen den Objekten aus zwei Listen anhand einer zu minimierenden Funktion (auch Distanz, Abstand, Ähnlichkeit). Die zwei Listen sind in dieser Arbeit die  $N_z$  Objektbeobachtungen und die  $N_x$  im System vorhandenen, prädierten Objekthypothesen zu einem bestimmten Zeitpunkt<sup>6</sup> (vgl. Fusionsprozess in Abbildung 7-2 auf Seite 68).

Für die Assoziation werden die Daten in denselben Koordinatenraum transformiert (Alignment, in dieser Arbeit den Objektraum, wie in Abschnitt 3.1 diskutiert), die optimale gemeinsame Referenz  $r_{assoc}$  bestimmt und die entsprechenden Referenzpunkte sowie Fehlerkovarianzen berechnet (bereits in letztem Abschnitt 7.3 vorgestellt).

Das Gating ist ein zusätzlicher Mechanismus zur Vermeidung von falschen Assoziationen. Dabei werden Assoziationen zwischen Objekthypothesen und Beobachtungen ausgeschlossen, die in eine physikalisch oder statistisch unwahrscheinliche Bewegung resultieren würden.

<sup>6</sup>Andere Möglichkeiten sind die Assoziation von Beobachtungen aus zwei Sensoren oder aus zwei Zeitpunkten.

### 7.4.1 Distanz zwischen zwei Objekten

Als Distanz (bzw. zu minimierende Funktion) zwischen einer Objektbeobachtung  $\mathbf{z}_s$  und einer prädizierten Objekthypothesenbeobachtung  $\hat{\mathbf{z}}$  wird eine Metrik  $d$  ausgewählt. Die Bedingungen, die eine Metrik erfüllen muss, sind:

1. Symmetrie:  $d(\mathbf{a}, \mathbf{b}) = d(\mathbf{b}, \mathbf{a}) \geq 0$
2. Dreiecksungleichung:  $d(\mathbf{a}, \mathbf{b}) \leq d(\mathbf{a}, \mathbf{c}) + d(\mathbf{c}, \mathbf{b})$
3. Nichtunterscheidbarkeit von Identitäten:  $d(\mathbf{a}, \mathbf{b}) = 0 \Leftrightarrow \mathbf{a} = \mathbf{b}$

Häufig wird der geometrische Abstand verwendet. Dabei ist lediglich die Position der Objekte relevant. Nach dem euklidischen Abstand kommen andererseits alle Komponenten in Betracht, es werden aber die Maßstäbe der unterschiedlichen Komponenten nicht berücksichtigt<sup>7</sup>.

$$d_E(\mathbf{z}_s, \hat{\mathbf{z}}) = \sqrt{(\mathbf{z}_s - \hat{\mathbf{z}})^T \cdot (\mathbf{z}_s - \hat{\mathbf{z}})} \quad (7.15)$$

Allgemeiner kann die Mahalanobis-Distanz<sup>8</sup> [98] verwendet werden. Diese bezieht sich auf zwei beliebige stochastische Vektoren, die mit der inversen Kovarianzmatrix gewichtet (normiert) werden, und damit die Distanz dimensionsfrei ist.

$$d_M(\mathbf{z}_s, \hat{\mathbf{z}}) = \sqrt{(\mathbf{z}_s - \hat{\mathbf{z}})^T \cdot \mathbf{S}^{-1} \cdot (\mathbf{z}_s - \hat{\mathbf{z}})} \quad (7.16)$$

Darin ist  $(\mathbf{z}_s - \hat{\mathbf{z}})$  die Innovation, die durch die Beobachtung  $\mathbf{z}_s$  in den Filter eingehen würde. Die Innovation wird mit der sogenannten Residual- oder Innovationskovarianzmatrix  $\mathbf{S}$  gewichtet. Diese ergibt sich als Summe der prädizierten und der beobachteten Kovarianzmatrizen:  $\mathbf{S} = \mathbf{P}_{\hat{\mathbf{z}}} + \mathbf{R}_s$  (siehe Abschnitt 7.2). Der Ausdruck  $d_M^2(\mathbf{z}_s, \hat{\mathbf{z}})$  weist eine  $\chi^2$ -Verteilung mit  $n_z$  Freiheitsgraden auf, wobei  $n_z$  die Dimension des Messvektors ist. Zur Verdeutlichung dieser Metrik sind in Abbildung 7-4.b Ellipsen mit gleicher Distanz zu den jeweiligen Beobachtungen aufgetragen. Aus einer unsicheren Beobachtung sind die Distanzschritte größer als aus einer sicheren.

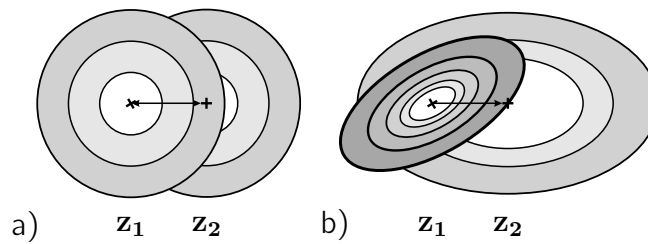
Weitere Möglichkeiten für die Distanzfunktion sind die normierte Mahalanobis-Distanz (vgl. [35]), die Bhattacharyya-Distanz oder die Korrelationskoeffizienten (vgl. [84]).

Für diese Arbeit wurde die Mahalanobis-Distanz als Assoziationsmetrik ausgewählt, denn sie ist einheitsfrei, nimmt die Unsicherheit des Objekts und der Beobachtung in Betracht und ihre statistische Verteilung ist bekannt  $\chi_{n_z}^2$ . Da die Assoziation für eine festgelegte logische Referenz durchgeführt wird (siehe letzten Abschnitt), wird hier die Mahalanobis-Distanz lediglich auf die Positionskomponenten  $x$  und  $y$  angewendet.

Die Mahalanobis-Distanz ist in zwei Situationen klein. Einerseits, wenn die Vektoren sehr nah aneinander liegen,  $(\mathbf{z}_s - \hat{\mathbf{z}}) \rightarrow 0$ , und andererseits wenn die Kovarianz (Ungenauigkeit) sehr groß ist,  $\mathbf{S} \rightarrow \infty$ . Letzteres ergibt sich, wenn eine Objekthypothese nach mehreren Messzyklen nicht

<sup>7</sup>Beispielsweise bei unterschiedlichen Einheiten von Position und Ausrichtung.

<sup>8</sup>Auch normierter Innovationsquadrat oder *normalized innovation square* (NIS) genannt.



**Abbildung 7-4:** Veranschaulichung der euklidischen (a) und der Mahalanobis-Distanz (b) zu zwei Zufallsgrößen. Die Ellipsen stellen Punkte mit gleicher Distanz zu jeweils  $\mathbf{z}_1$  und  $\mathbf{z}_2$  dar. Je größer die Distanz, desto dunkler sind sie schattiert. In b) hat  $\mathbf{z}_1$  eine kleinere Unsicherheit. Nach der Metrik  $d_{M'}(\mathbf{z}_1, \mathbf{z}_2)$  würde die Mahalanobis-Distanz von  $\mathbf{z}_1$  ausgehend ausgewählt.

beobachtet wurde, womit sich seine Unsicherheit erhöht hat. Damit besitzt diese kleine Distanzen zu allen Beobachtungen, was die Assoziation verfälscht. Um diesen Effekt zu vermeiden, schlägt Stüker in [144] vor, einen Strafterm zur Distanz zu addieren, was jedoch negative Folgen im Gating hat. In dieser Arbeit wird die Mahalanobis-Distanz so angepasst, dass die Unsicherheiten aus der prädizierten Objekthypothese und die aus der Beobachtung getrennt voneinander betrachtet werden (asymmetrisch), und nur die GröÙte der beiden Distanzen ausgewählt wird, also aus dem sichereren Objekt (siehe Pfeil in Abbildung 7-4.b):

$$d_{M'}(\mathbf{z}_s, \hat{\mathbf{z}}) = \max \left( \sqrt{(\mathbf{z}_s - \hat{\mathbf{z}})^T \cdot \mathbf{P}_{\hat{\mathbf{z}}}^{-1} \cdot (\mathbf{z}_s - \hat{\mathbf{z}})}, \sqrt{(\mathbf{z}_s - \hat{\mathbf{z}})^T \cdot \mathbf{R}_s^{-1} \cdot (\mathbf{z}_s - \hat{\mathbf{z}})} \right) \quad (7.17)$$

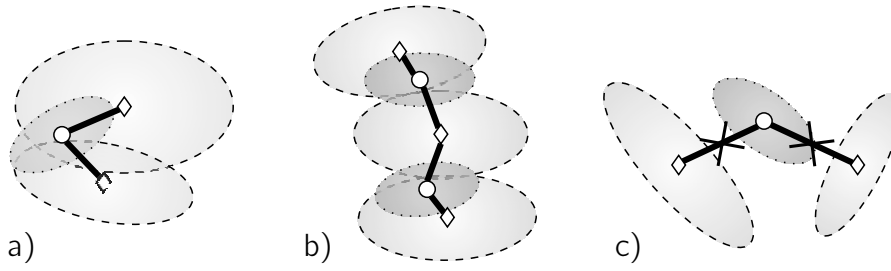
Damit wird der zuvor erwähnte Effekt vermieden. Es kann dabei bewiesen werden, dass diese Funktion ebenso die Eigenschaften einer Metrik besitzt<sup>9</sup>.

## 7.4.2 Zuweisung

Die eigentliche Assoziation besteht in der Zuweisung zwischen Objektbeobachtungen und Objekthypothesen.

In Fahrsituationen ist die Anzahl von Objekten in der Regel gering und es treten wenige Gesamtbeobachtungsfehler (beispielsweise durch Ruckeln des Fahrzeugs) auf. Daher ist ein *Nearest-Neighbor(s)*-Algorithmus ausreichend und wird hier verwendet. Um frühe Entscheidungen zu vermeiden, werden dabei mehrere Beobachtungen pro Objekthypothese und auch mehrere Objekthypothesen pro Beobachtung erlaubt, solange diese nicht durch das Gating ausgeschlossen werden – siehe nächsten Unterabschnitt und Abbildung 7-5. Dies ist dadurch begründet, weil sowohl die Kamera als auch der Laserscanner zum Teil mehrere Beobachtungen aus jedem Objekt erzeugen. Die Einbringung mehrerer Beobachtungen (aus einem oder mehreren Sensoren) ist durch den Informationsfilter berücksichtigt, wie in Abschnitt 7.5 eingeführt wird. Falls falsche

<sup>9</sup>Das Maximum zweier Metriken besitzt ebenso die Eigenschaften einer Metrik.



**Abbildung 7-5:** Objektassoziation und Gating-Strategie. Objekthypothesen sind durch Kreise, Beobachtungen durch Rauten und Unsicherheiten durch Ellipsen dargestellt. a) Mehrere Beobachtungen für eine Objekthypothese; b) die Mittlere Beobachtung ist mit beiden Objekthypothesen assoziiert; c) durch das Gating ausgeschlossene Assoziationen.

Zuweisungen entstehen, werden diese in weiteren Assoziationsschritten wie das Gating oder die Objektverwaltung wieder aufgehoben.

### 7.4.3 Gating

Ein wesentlicher Schritt zur Vermeidung von Fehlzuordnungen ist das Gating. Die grundsätzliche Vorgehensweise beim Gating besteht darin, einen Suchbereich um jede prädizierte Objekthypothese zu definieren, in dem sich Beobachtungen für das Objekt befinden können.

Die in dieser Arbeit durchgeführte Gating-Strategie besteht aus dem Vergleich der Assoziationsdistanz mit einem Schwellwert. Um den Schwellwert in Bezug auf die angepasste Mahalanobis-Distanz ( $d_{M'}$ ) festzulegen, wird deren bekannte  $\chi^2$ -Verteilung betrachtet und ein bestimmtes Quantil  $\chi^2_{n_z, 1-\alpha}$  ausgewählt, wie beispielsweise für  $\alpha = 0,05$ .

$$d_{M'}^2(\mathbf{z}_s, \hat{\mathbf{z}}) \leq \chi^2_{n_z, 1-\alpha} \quad (7.18)$$

Durch die Verwendung der asymmetrischen Mahalanobis-Distanz entspricht das Gating keinem konkreten Bereich um das Objekt, sondern entsteht *ad hoc* für jede Kombination von Beobachtung und Objekthypothese automatisch.

Es ist an dieser Stelle zu betonen, dass die Priorität des Gatings auf der Vermeidung von Fehlzuordnungen bei möglichst keinem Verfehlen von wirklichen Zuordnungen liegt. Da diese Entscheidung aufgrund von Mehrdeutigkeiten oft nicht erreicht wird, wird in dieser Arbeit neben den beiden Möglichkeiten *assoziiert* und *nichtassoziiert* eine dritte für *unklare* Zuordnungen hinzugefügt, die unwahrscheinliche aber mögliche Zuordnungen darstellt. Dies wird durch einen zweiten Schwellwert erreicht, zum Beispiel mit  $\alpha' = 2\alpha$ . Diese Hysterese wird im Abschnitt 7.8 im Zusammenhang mit den Objektverwaltungsaufgaben aufgegriffen.

## 7.5 Schätzung mittels des Informationsfilters

Der Informationsfilter (Maybeck [106], Rao et al. [117], Brown und Hwang [23] sowie Mutambara [109]) stellt eine mathematische Umformulierung des Kalman-Filters dar, in der die Information (im Sinne der Fisher-Matrix) über den Objektzustand eher als seine Kovarianzen betrachtet wird

(siehe Anhang C für mehr Details). Die Vorteile dieses Filters sind die Eignung für die Fusion von mehreren gleichzeitigen Beobachtungen, die direkte Objektinitialisierung und der geringe Rechenaufwand. Der Filter bietet eine Multi-Sensor-Fusion für sich, was im Kalman-Filter nur über Modell- oder Filteranpassungen möglich ist [13, 43, 89, 95, 167, 35]. Ein Partikelfilter, wie er bspw. in [124] verwendet wird, wurde aus Echtzeitgründen ausgeschlossen.

Es wird ein asynchroner Informationsfilter für jede Objekthypothese initialisiert. In jedem Zyklus wird dann jede im System vorhandene Objekthypothese mit den dazu assoziierten Beobachtungen  $\mathbf{z}_s(t_i)$  aktualisiert bzw. neu geschätzt, mit Sensor  $s \in \{cam, ls\}$ . Dabei ist es unwesentlich, ob diese aus einem oder mehreren Sensoren stammen. Der Informationsfilter verwendet statt dem Zustandsvektor  $\mathbf{x}(t_i)$  und der Kovarianzmatrix  $\mathbf{P}(t_i)$  die Informationsmatrix  $\mathbf{Y}(t_i)$  und den Informationszustandsvektor  $\mathbf{y}(t_i)$ , für  $t_i$  einem beliebigen Zeitpunkt:

$$\mathbf{Y}(t_i|t_j) \triangleq \mathbf{P}^{-1}(t_i|t_j) \quad (7.19)$$

$$\hat{\mathbf{y}}(t_i|t_j) \triangleq \mathbf{P}^{-1}(t_i|t_j) \cdot \hat{\mathbf{x}}(t_i|t_j) \quad (7.20)$$

### 7.5.1 Prädiktion

Der Schritt der Prädiktion entspricht den in Abschnitt 7.2 gegebenen Gleichungen (7.6) und (7.7) für  $\hat{\mathbf{x}}(t_k|t_{k-1})$  und  $\mathbf{P}(t_k|t_{k-1})$  aus den letzten Schätzung  $\hat{\mathbf{x}}(t_{k-1}|t_{k-1})$  und  $\mathbf{P}(t_{k-1}|t_{k-1})$ , der Übergangsfunktion  $\mathbf{F}(t_k)$  und dem Prozessrauschen  $\mathbf{Q}(t_k)$ . Anschließend werden auch die prädierten Informationsstrukturen  $\mathbf{Y}(t_k|t_{k-1})$  und  $\hat{\mathbf{y}}(t_k|t_{k-1})$  anhand der Gleichungen (7.19) bzw. (7.20) berechnet. Dazu wird die Prädiktion in den Beobachtungsraum transformiert,  $\hat{\mathbf{z}}(t_k|t_{k-1})$  und  $\mathbf{P}_{\hat{\mathbf{z}}}(t_k|t_{k-1})$  (Gleichungen (7.8) und (7.9)).

### 7.5.2 Schätzung

Nach Empfang und Zuordnung (Abschnitt 7.4) eines oder mehrerer Messvektoren  $\mathbf{z}_s$  mit dem Messrauschen  $\mathbf{R}_s$  wird deren jeweiliger Informationsbeitrag wie folgt definiert:

$$\mathbf{I}_s(t_i) \triangleq \mathbf{H}_s^T(t_i) \cdot \mathbf{R}_s^{-1}(t_i) \cdot \mathbf{H}_s(t_i) \quad (7.21)$$

$$\mathbf{i}_s(t_i) \triangleq \mathbf{H}_s^T(t_i) \cdot \mathbf{R}_s^{-1}(t_i) \cdot \mathbf{z}_s(t_i) \quad (7.22)$$

Damit wird die Schätzung der neuen Informationsmatrix und des neuen Informationszustandes der betrachteten Objekthypothese als einfache Summe über die Beiträge aller zu fusionierenden

Beobachtungen im Zeitpunkt  $t_k$  durchgeführt:

$$\mathbf{Y}(t_k|t_k) = \mathbf{Y}(t_k|t_{k-1}) + \sum_s \mathbf{I}_s(t_k) \quad (7.23)$$

$$\hat{\mathbf{y}}(t_k|t_k) = \hat{\mathbf{y}}(t_k|t_{k-1}) + \sum_s \mathbf{i}_s(t_k) \quad (7.24)$$

Wurde eine Objekthypothese in diesem Messzyklus nicht beobachtet (assoziiert), so ist die Summe der Beiträge leer. Die Prädiktion wird in diesem Fall als neue Schätzung ausgegeben. Für die darauf folgende Prädiktion wird jedoch die letzte Schätzung verwendet, welche mit Beiträgen versehen wurde.

Wurden mehrere Beobachtungen aus einem Sensor zu einer bestimmten Schätzung assoziiert, so lässt sich auf gesplittete Sensordaten schließen. Diese liefern jedoch nur so viel Information zur Schätzung der Position wie eine einzige Beobachtung. Daher werden deren Beiträge in solchen Fällen gewichtet, so dass sie die Schätzung nicht verfälschen.

### 7.5.3 Filtereigenschaften

Die Vorteile des Informationsfilters im Vergleich zum Kalman-Filter sind nach Mutambara [109] im Wesentlichen folgende:

- Einfache Verteilbarkeit (deswegen auch genannt *decentralized Kalman filter*) bzw. Eignung für Multi-Sensor-Systeme, in denen mehrere Beobachtungen gleichzeitig vorkommen können<sup>10</sup>;
- Direkte Initialisierung: wird ein Objekt zum ersten Mal beobachtet, so werden die Informationsmatrix und der Informationszustand auf Null initialisiert:

$$\mathbf{Y}(t_0|t_0) = \mathbf{0}; \hat{\mathbf{y}}(t_0|t_0) = \mathbf{0}$$

Aus praktischen Gründen, um die Invertierbarkeit von  $\mathbf{Y}(t_0|t_0)$  zu gewährleisten, werden deren Diagonalelemente mit sehr kleinen Werten initialisiert.

- Wenn der Zustandsraum weniger Dimensionen hat als der Beobachtungsraum, sind die zu invertierenden Matrizen kleiner als beim Kalman-Filter, und damit rechnerisch weniger aufwändig.

Die Nachteile sind:

- Geringe Intuitivität von  $\mathbf{Y}(t_i|t_j)$  und  $\hat{\mathbf{y}}(t_i|t_j)$ , wobei es jederzeit möglich ist, die entsprechenden  $\mathbf{x}(t_i|t_j)$  und  $\mathbf{P}(t_i|t_j)$  anhand von (7.19) und (7.20) zu berechnen.
- Wenn der Zustandsraum mehr Dimensionen hat als der Beobachtungsraum, dann sind die zu invertierenden Matrizen größer als beim Kalman-Filter, und damit rechnerisch aufwändiger. In dieser Arbeit haben beide Räume die gleiche Dimensionalität.

<sup>10</sup>Der Filter ist sowohl für das Tracking aus einem Sensor als auch für die konkurrierende Fusion mehrerer Sensoren geeignet.

Der Informationsfilter kann ebenso für nichtlineare Systeme angepasst werden. Das ist der sogenannte erweiterte Informationsfilter (EIF, [109]).

## 7.6 Partielle Beobachtbarkeit der Objekte

Wie in früheren Abschnitten erläutert wurde, verändert sich das Aussehen eines Objektes je nach Situation und Sensor zum Teil stark. Dabei können sich die optimale Referenz des Objekts (Abschnitt 7.3) sowie die Beobachtbarkeit dessen einzelner Komponenten<sup>11</sup> verändern. Die Beobachtbarkeit bezieht sich sowohl auf anschauliche Komponenten (wie Breite und Länge eines Objekts) als auch auf abgeleitete oder von mehreren Messungen abhängige Komponenten (wie Geschwindigkeit, oder der Mittelpunkt eines Objekts). Konkrete Beispiele werden in Anhang D vorgestellt. In diesem Abschnitt werden die bisher vorgestellten Überlegungen mit weiteren Betrachtungen der partiellen Beobachtbarkeit der verfolgten Objekthypothesen ergänzt.

Stüker beschäftigt sich in [144] formell mit dieser Problematik. Er setzt dabei auf die Definitionen von vollständiger bzw. lokaler Beobachtbarkeit auf (vgl. Chen [32] und Bar-Shalom [10]). In der Kontrolltheorie wird ein System als vollständig beobachtbar (*observable*) definiert, wenn dessen aktueller Zustand für beliebige Zustands- und Steuerungsvektoren anhand einer endlichen Anzahl von Beobachtungen bestimmt werden kann. Die Beobachtbarkeit eines zeitinvarianten linearen Systems kann anhand des Rangs folgender Matrix bestimmt werden:

$$\mathbf{M}(t_i) = \begin{bmatrix} \mathbf{H}_s(t_i) \\ \mathbf{H}_s(t_i) \cdot \mathbf{F}(t_i) \\ \mathbf{H}_s(t_i) \cdot \mathbf{F}^2(t_i) \\ \vdots \\ \mathbf{H}_s(t_i) \cdot \mathbf{F}^{n_x-1}(t_i) \end{bmatrix} \quad (7.25)$$

Besitzt diese Matrix den vollen Rang  $n_x$ , so ist das System vollständig beobachtbar. Durch diesen Aufbau sind nicht nur die direkt beobachtbaren Komponenten (wie Position oder Größe) sondern auch die davon abgeleiteten (wie Geschwindigkeit und Beschleunigung) betrachtet.

Diese Betrachtung wird in dieser Arbeit erweitert, um die Beobachtbarkeit jeder einzelnen Systemkomponente  $x_k$  zu prüfen. Dies bedeutet, den Rang von  $\mathbf{M}(t_i)$  mit dem Rang von  $\mathbf{M}^k(t_i)$ , welche die Matrix  $\mathbf{M}(t_i)$  ohne die  $k$ -te Spalte darstellt, zu vergleichen.

$$\mathbf{M}^k(t_i); m_{ij}^k = \begin{cases} m_{ij} & \text{wenn } j \neq k, \\ 0 & \text{sonst.} \end{cases} \quad (7.26)$$

Die  $k$ -te Komponente des Systems ist beobachtbar, wenn sich der Rang der Matrix  $\mathbf{M}(t_i)$  durch Hinzufügen dieser Spalte erhöht, also

<sup>11</sup>Die Beobachtbarkeit einer Komponente sagt nichts über deren Genauigkeit oder Qualität aus.

$$x_k \text{ beobachtbar} \leftrightarrow \text{Rang}(\mathbf{M}(t_i)) - \text{Rang}(\mathbf{M}^k(t_i)) = 1. \quad (7.27)$$

Die Beobachtbarkeit der einzelnen Komponenten eines Mess- oder Zustandsvektors wird hier in Form der in Abschnitt 3.2 auf Seite 33 eingeführten Beobachtbarkeitsmatrix  $\mathbf{V}_{\hat{\mathbf{x}}}$  bzw.  $\mathbf{V}_{\mathbf{z}_s}$  dargestellt (ähnlich wie der von Stüker in [144] vorgestellte Initialisierungsvektor). Die Diagonalelemente dieser Matrix, die beobachtbare Komponenten darstellen, haben den Wert Eins, die restlichen betragen Null. Hier werden die Ausdrücke aus diesem Kapitel (Prädiktion, Assoziation, Informationsfilter usw.) anhand der Beobachtbarkeitsmatrix angepasst, so dass nur beobachtbare Zustandskomponenten vom Filter verfolgt werden.

Die Beobachtbarkeitsmatrix der eingehenden Objektbeobachtungen  $\mathbf{V}_{\mathbf{z}_s}$  wird anhand der direkten Beobachtbarkeitsüberprüfungen der Komponenten gefüllt, wie in den Abschnitten 5.5 und 6.2 beschrieben. Die Beobachtbarkeitsmatrix der verfolgten Objekthypothesen muss mitgeschätzt werden. Dafür wird sie mit Null initialisiert (keine Komponente ist beobachtbar), und wird anhand der Beobachtbarkeit der eingehenden Objektbeobachtungen in jedem Zyklus aktualisiert. Eine bereits beobachtbare Komponente des Objektzustands wird immer als beobachtbar beibehalten. Wird diese Komponente nicht mehr durch die Sensoren erfasst, so wird sie trotzdem prädiziert und eine immer höhere Unsicherheit besitzen. Eine nichtbeobachtbare Komponente (und auch die entsprechenden Kovarianzen) wird stets mit Null initialisiert.

### 7.6.1 Prädiktion des Zustandsvektors

Bei der Prädiktion der Objekthypothesen sollen ausschließlich die beobachtbaren Komponenten ermittelt werden. Dafür werden die Gleichungen (7.6) und (7.7) von Seite 71 wie folgt angepasst.

$$\hat{\mathbf{x}}(t_k|t_{k-1}) = \mathbf{F}(t_k)\hat{\mathbf{x}}(t_{k-1}|t_{k-1}) \quad (7.28)$$

$$\mathbf{P}(t_k|t_{k-1}) = \mathbf{F}(t_k)\mathbf{P}(t_{k-1}|t_{k-1})\mathbf{F}^T(t_k) + \mathbf{V}_{\hat{\mathbf{x}}}(t_{k-1}) \cdot \mathbf{Q}(t_k) \cdot \mathbf{V}_{\hat{\mathbf{x}}}(t_{k-1}) \quad (7.29)$$

Da alle unbeobachtbaren Komponenten mit Null initialisiert wurden, werden diese in der Prädiktion  $\hat{\mathbf{x}}$  ebenfalls Null betragen und können direkt mit  $\mathbf{F}$  multipliziert werden. Die Kovarianzen der unbeobachtbaren Komponenten sind ebenfalls mit Null initialisiert und sollen den Wert Null behalten. Damit das Prozessrauschen dieser Komponenten keinen Einfluss hat, wird die Matrix  $\mathbf{Q}$  mit der Beobachtbarkeitsmatrix des Objekts  $\mathbf{V}_{\hat{\mathbf{x}}}$  multipliziert.

Die restlichen sowohl direkt als auch indirekt beobachtbaren Komponenten und deren Kovarianzen werden wie im Standardfilter prädiziert.

Die Beobachtbarkeitsmatrix der Prädiktion ändert sich nicht:  $\mathbf{V}_{\hat{\mathbf{x}}}(t_k|t_{k-1}) = \mathbf{V}_{\hat{\mathbf{x}}}(t_{k-1})$ .



## 7.6.2 Prädiktion des Messvektors

Als Vorschritt zur Assoziation werden die prädizierten Objekthypothesen in den Beobachtungsraum transformiert. Diese Transformation ist in dieser Arbeit nicht notwendig, da der Zustandsraum und der Beobachtungsraum der gleiche sind (siehe Abschnitt 3.1). Der Vollständigkeit halber werden hier die Ausdrücke (7.8) und (7.9) von Seite 71 auf die partielle Beobachtbarkeit angepasst.

Im Umgang mit der partiellen Beobachtbarkeit können Situationen auftreten, in denen nur ein Teil der Zustandskomponenten, von denen  $\hat{\mathbf{z}}$  abhängt, beobachtbar sind (siehe Beispiel 3 in Anhang D). Daher ist es notwendig die Beobachtbarkeit  $v_{\hat{\mathbf{z}}}$  der einzelnen Komponenten von  $\hat{\mathbf{z}}$  aus der Beobachtbarkeit von  $\hat{\mathbf{x}}$  zu prüfen.

Die für jede Beobachtungskomponente  $z_{s,i}$  des Sensors  $s$  nötigen Komponenten  $x_j$  des Zustandsvektors lassen sich anhand der Beobachtungsmatrix  $\mathbf{H}_s$  wie folgt ausdrücken:

$$\mathbf{D}_s; d_{ij} = \begin{cases} 1 & \text{wenn } h_{s,ij} \neq 0, \\ 0 & \text{sonst.} \end{cases} \quad (7.30)$$

Wenn  $d_{s,ij} = 1$ , ist weil  $x_j$  einen Einfluss auf  $z_{s,i}$  für Sensor  $s$  hat.

Eine Komponente des prädizierten Messvektors ist nur dann beobachtbar, wenn sie für diesen Sensor bei voller Zustandsbeobachtbarkeit beobachtbar ist und alle nötigen Zustandsvektorkomponenten beobachtbar sind, also:

$$v_{\hat{\mathbf{z}}_{s,i}} = \begin{cases} 1 & \text{wenn } \sum_j d_{s,ij} > 0 \cap \sum_j d_{s,ij} = \sum_j d_{s,ij} v_{\hat{\mathbf{x}}_j}, \\ 0 & \text{sonst.} \end{cases} \quad (7.31)$$

Es wäre nicht richtig,  $h_{s,ij}$  statt  $d_{s,ij}$  zu verwenden, da sich mehrere Komponenten der Spalte  $j$  aufheben könnten, womit  $\sum_j h_{s,ij} = 0$  und die Komponente  $i$  fälschlicherweise als unbeobachtbar eingestuft werden würde ( $\sum_j d_{s,ij} > 1$ ).

Die Gleichungen (7.8) und (7.9) auf Seite 71 zur Prädiktion des Messvektors werden nun mit der errechneten Beobachtbarkeitsmatrix  $\mathbf{V}_{\hat{\mathbf{z}}_s}$  ergänzt:

$$\hat{\mathbf{z}}_s(t_k|t_{k-1}) = \mathbf{V}_{\hat{\mathbf{z}}_s}(t_k|t_{k-1}) \cdot \mathbf{H}_s(t_k) \cdot \hat{\mathbf{x}}(t_k|t_{k-1}) \quad (7.32)$$

$$\mathbf{P}_{\hat{\mathbf{z}}_s}(t_k|t_{k-1}) = \mathbf{V}_{\hat{\mathbf{z}}_s}(t_k|t_{k-1}) \cdot \mathbf{H}_s(t_k) \cdot \mathbf{P}(t_k|t_{k-1}) \cdot \mathbf{H}_s^T(t_k) \cdot \mathbf{V}_{\hat{\mathbf{z}}_s}(t_k|t_{k-1}) \quad (7.33)$$

Damit werden nur beobachtbare Messkomponenten berechnet, während die nichtbeobachtbaren mit Null gefüllt werden (siehe dazu erläuternde Beispiele in Anhang D).

### 7.6.3 Assoziation

Um die eingehenden Beobachtungen mit den prädizierten Beobachtungen zu vergleichen, dürfen nur die gemeinsam beobachtbaren Komponenten berücksichtigt werden. So werden jeweils die Vektoren und Kovarianzmatrizen der einen Beobachtung mit der Beobachtbarkeitsmatrix der anderen Beobachtung multipliziert:

$$\hat{\mathbf{z}}_{\mathbf{s},\mathbf{v}}(t_k|t_{k-1}) = \mathbf{V}_{\mathbf{z}_{\mathbf{s}}}(t_k) \cdot \hat{\mathbf{z}}_{\mathbf{s}}(t_k|t_{k-1}) \quad (7.34)$$

$$\mathbf{z}_{\mathbf{s},\mathbf{v}}(t_k) = \mathbf{V}_{\hat{\mathbf{z}}_{\mathbf{s}}}(t_k|t_{k-1}) \cdot \mathbf{z}_{\mathbf{s}}(t_k) \quad (7.35)$$

$$\mathbf{P}_{\hat{\mathbf{z}}_{\mathbf{s}},\mathbf{v}}(t_k|t_{k-1}) = \mathbf{V}_{\mathbf{z}_{\mathbf{s}}}(t_k) \cdot \mathbf{P}_{\hat{\mathbf{z}}_{\mathbf{s}}}(t_k|t_{k-1}) \cdot \mathbf{V}_{\mathbf{z}_{\mathbf{s}}}(t_k) \quad (7.36)$$

$$\mathbf{P}_{\mathbf{z}_{\mathbf{s}},\mathbf{v}}(t_k|t_{k-1}) = \mathbf{V}_{\hat{\mathbf{z}}_{\mathbf{s}}}(t_k|t_{k-1}) \cdot \mathbf{R}_{\mathbf{s},\mathbf{v}}(t_k) \cdot \mathbf{V}_{\hat{\mathbf{z}}_{\mathbf{s}}}(t_k|t_{k-1}) \quad (7.37)$$

Hierbei beinhaltet  $\mathbf{R}_{\mathbf{s},\mathbf{v}}(t_k) = \mathbf{V}_{\mathbf{z}_{\mathbf{s}}}(t_k) \cdot \mathbf{R}_{\mathbf{s}}(t_k) \cdot \mathbf{V}_{\mathbf{z}_{\mathbf{s}}}(t_k)$  nur die in der Objektbeobachtung beobachteten Komponenten<sup>12</sup>.  $\mathbf{V}_{\mathbf{z}_{\mathbf{s}}}$  ist die mitgelieferte Beobachtbarkeit der zu assoziierenden Objektbeobachtung (siehe Kapiteln 5.5 und 6.1). Damit bleiben in der Objektbeobachtung und in der prädizierten Objektbeobachtung nur die gemeinsamen beobachtbaren Komponenten ungleich Null, so dass die Mahalanobis-Distanz – Ausdruck (7.17) auf Seite 75 – berechnet werden kann.

### 7.6.4 Schätzung der Zustandsbeobachtbarkeit

Anhand der Beobachtbarkeit der assoziierten Objektbeobachtungen wird nun die Beobachtbarkeit jeder Objekthypothese neu geschätzt bzw. gegebenenfalls erweitert.

Das Setzen einer Komponente der Objekthypothese auf beobachtbar wird anhand einer ähnlichen Prüfung wie Gleichung (7.27) durchgeführt. Dabei wird in Gleichung (7.25) die Beobachtungsmatrix  $\mathbf{H}_{\mathbf{s}}(t_i)$  mit der Beobachtbarkeit der eingehenden Objektbeobachtung  $\mathbf{V}_{\mathbf{z}_{\mathbf{s}}}(t_i)$  multipliziert, so dass nur die beobachtbaren Komponenten Einfluss haben:

$$\mathbf{M}_{\mathbf{V}}(t_i) = \begin{bmatrix} \mathbf{V}_{\mathbf{z}_{\mathbf{s}}}(t_i) \cdot \mathbf{H}_{\mathbf{s}}(t_i) \\ \mathbf{V}_{\mathbf{z}_{\mathbf{s}}}(t_i) \cdot \mathbf{H}_{\mathbf{s}}(t_i) \cdot \mathbf{F}(t_i) \\ \mathbf{V}_{\mathbf{z}_{\mathbf{s}}}(t_i) \cdot \mathbf{H}_{\mathbf{s}}(t_i) \cdot \mathbf{F}^2(t_i) \\ \vdots \\ \mathbf{V}_{\mathbf{z}_{\mathbf{s}}}(t_i) \cdot \mathbf{H}_{\mathbf{s}}(t_i) \cdot \mathbf{F}^{n_x-1}(t_i) \end{bmatrix} \quad (7.38)$$

Daraus ergibt sich die Beobachtbarkeit anhand der Beobachtung  $\mathbf{z}_{\mathbf{s}}$ :

<sup>12</sup>Die Sensoren liefern stets die gesamte Messrauschenmatrix  $\mathbf{R}_{\mathbf{s}}$ , unabhängig von der Beobachtbarkeit.

$$x_k \text{ beobachtbar} \leftrightarrow \text{Rang}(\mathbf{M}_{\mathbf{V}}(t_i)) - \text{Rang}(\mathbf{M}_{\mathbf{V}}^{\mathbf{k}}(t_i)) = 1. \quad (7.39)$$

Die aktualisierte Beobachtbarkeit der Objekthypothese wird als Gesamtmenge der prädizierten Beobachtbarkeit  $\mathbf{V}_{\hat{\mathbf{x}}}$  und der hinzugekommenen  $\mathbf{V}_{\mathbf{x}_s}$  aus Sensor  $s$  aktualisiert:

$$\mathbf{V}_{\hat{\mathbf{x}}}(t_k|t_k); v_{\hat{x},ii}(t_k|t_k) = \max(v_{\hat{x},ii}(t_k|t_{k-1}), v_{x_s,ii}(t_k)) \quad (7.40)$$

In Anhang D sind einige Beispiele vorgestellt, die die partielle Beobachtbarkeit der Objekthypothesen anhand unterschiedlicher Beobachtungen herleitet.

### 7.6.5 Zustandsschätzung

Der Informationsfilter muss wie folgt aus Gleichung (7.19) auf Seite 77 angepasst werden. Die Informationsstrukturen werden leicht anders definiert, so dass der beobachtbare Teil von  $\mathbf{P}$ , also die reguläre Teilmatrix, invertiert werden kann,

$$\mathbf{Y}_{\mathbf{v}}(t_i|t_j) \triangleq [\mathbf{V}_{\hat{\mathbf{x}}}(t_i|t_j) \cdot \mathbf{P}(t_i|t_j) \cdot \mathbf{V}_{\hat{\mathbf{x}}}(t_i|t_j) + \overline{\mathbf{V}}_{\hat{\mathbf{x}}}(t_i|t_j)]^{-1} - \overline{\mathbf{V}}_{\hat{\mathbf{x}}}(t_i|t_j), \quad (7.41)$$

wobei  $\overline{\mathbf{V}}_{\hat{\mathbf{x}}}(t_i|t_j) = \mathbf{I} - \mathbf{V}_{\hat{\mathbf{x}}}(t_i|t_j)$  und  $\mathbf{I}$  die Einheitsmatrix darstellen. Dabei wird vor der Matrixinversion eine Eins an jede bisher unbeobachtbare Diagonalstelle von  $\mathbf{P}$  gesetzt. Nach der Inversion wird sie wieder auf Null gesetzt<sup>13</sup>. Für den Informationszustandsvektor aus Gleichung (7.20) wird:

$$\hat{\mathbf{y}}_{\mathbf{v}}(t_i|t_j) \triangleq \mathbf{Y}_{\mathbf{v}}(t_i|t_j) \cdot \hat{\mathbf{x}}(t_i|t_j) \quad (7.42)$$

Zunächst wird der Informationsbeitrag berechnet. Dabei wird die Beobachtbarkeit der Objektbeobachtung  $\mathbf{V}_{\mathbf{z}_s}$  betrachtet, einmal in den Zustandsraum  $\mathbf{V}_{\mathbf{x}_s}$  und einmal zurück in den Beobachtungsraum als  $\mathbf{V}_{\mathbf{z}'_s}$  transformiert. Dazu werden jeweils die Gleichungen (7.38) und (7.40) und die Gleichungen (7.30) und (7.31) verwendet. Das ist ein wichtiger Schritt, denn wenn eine Zustandskomponente von zwei Beobachtungskomponenten abhängt und nur eine davon beobachtbar ist, so soll sie keinen Einfluss auf die Zustandskomponente haben (siehe Beispiel 3 im Anhang D). Statt den Gleichungen (7.21) und (7.22) werden folgende verwendet,

$$\mathbf{I}_{\mathbf{s},\mathbf{v}}(t_i) \triangleq \mathbf{H}_{\mathbf{s}}^T(t_i) \left[ (\mathbf{V}_{\mathbf{z}'_s}(t_i) \mathbf{R}_{\mathbf{s}}(t_i) \mathbf{V}_{\mathbf{z}'_s}(t_i) + \overline{\mathbf{V}}_{\mathbf{z}'_s}(t_i))^{-1} - \overline{\mathbf{V}}_{\mathbf{z}'_s}(t_i) \right] \mathbf{H}_{\mathbf{s}}(t_i) \quad (7.43)$$

$$\mathbf{i}_{\mathbf{s},\mathbf{v}}(t_i) \triangleq \mathbf{H}_{\mathbf{s}}^T(t_i) \left[ (\mathbf{V}_{\mathbf{z}'_s}(t_i) \mathbf{R}_{\mathbf{s}}(t_i) \mathbf{V}_{\mathbf{z}'_s}(t_i) + \overline{\mathbf{V}}_{\mathbf{z}'_s}(t_i))^{-1} - \overline{\mathbf{V}}_{\mathbf{z}'_s}(t_i) \right] \mathbf{z}_{\mathbf{s}}(t_i) \quad (7.44)$$

<sup>13</sup>Um zu beweisen, dass die reguläre Teilmatrix von  $\mathbf{Y}_{\mathbf{v}}$  der Inverse des beobachtbaren Teils von  $\mathbf{P}$  entspricht, werden die Adjunkten und der Laplacesche Entwicklungssatz verwendet.

mit  $\bar{\mathbf{V}}_{\mathbf{z}_s'}(t_i) = \mathbf{I} - \mathbf{V}_{\mathbf{z}_s'}(t_i)$ . Da in dieser Arbeit der Zustandsraum dem Beobachtungsraum entspricht, fallen in diesen Ausdrücken die Multiplikationen mit  $\mathbf{H}_s$  und  $\mathbf{H}_s^T$  aus. Die Matrix  $\mathbf{I}_{s,v}$  und der Vektor  $\mathbf{i}_{s,v}$  beinhalten nur Information über die Zustandskomponenten, die sich aus der beobachtbaren Beobachtungskomponenten herleiten, während die restlichen Komponenten Null betragen.

Die eigentliche Schätzung entspricht:

$$\mathbf{Y}_v(t_k|t_k) = \mathbf{Y}_v(t_k|t_{k-1}) + \sum_s \mathbf{I}_{s,v}(t_k) \quad (7.45)$$

$$\hat{\mathbf{y}}_v(t_k|t_k) = \hat{\mathbf{y}}_v(t_k|t_{k-1}) + \sum_s \mathbf{i}_{s,v}(t_k) \quad (7.46)$$

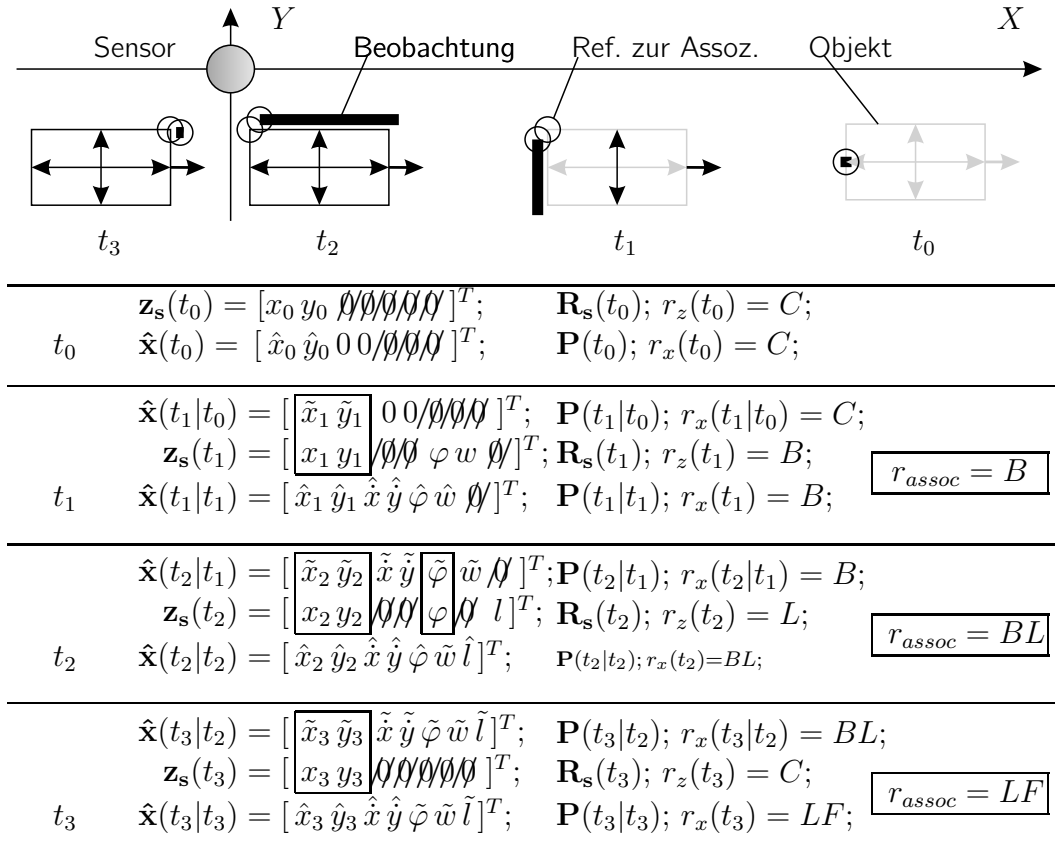
Dies kann wiederum in den Zustandsraum anhand der Gleichungen (7.41) und (7.42) mithilfe der geschätzten Beobachtbarkeit  $\mathbf{V}_{\hat{\mathbf{x}}}(t_k|t_k)$  umgewandelt werden.

Falls neue Komponenten durch diese Beobachtung beobachtbar werden, werden diese ihren Beitrag in  $\mathbf{I}_{s,v}$  und  $\mathbf{i}_{s,v}$  haben, womit sie auch initialisiert werden. Mit diesen Ausdrücken können gleichzeitige Beobachtungen deren Beitrag zu gleichen und unterschiedlichen Komponenten eines Objekts leisten, womit die Fusion konkurrierende und komplementäre Anteile besitzt. Allerdings ist mit diesen Methoden keine kooperative Fusion möglich: Benötigt die Berechnung einer Zustandskomponente die Beobachtbarkeit von zwei Beobachtungskomponenten (wie die Breite eines Objekts aus einer linken und einer rechten Position), die von unterschiedlichen Sensoren geliefert werden, so werden sie durch die Ausdrücke (7.43) und (7.44) jeweils keinen Einfluss auf den Zustand haben.

### 7.6.6 Zusammenfassung der Beobachtbarkeit

Mit den vorgestellten Änderungen der Filtergleichungen ist es möglich, die Beobachtbarkeit der Objekte im Laufe der Zeit mitzuverfolgen, die sich aufgrund der Perspektive sowie von Verdeckungen zwischen Objekten verändert. Die Komplexität und der Rechenaufwand des Algorithmus sind dabei nur gering erhöht worden. Die Effekte dieser Betrachtungen können anhand eines Beispiels mit sich ändernder Referenz und Beobachtbarkeit in Abbildung 7-6 gesehen werden. Dabei wird ein Objekt zunächst in großer Entfernung erfasst ( $t_0$ ), wodurch alle Komponenten bis auf die Position unbeobachtbar sind. Später in  $t_1$  werden seine Breite, Ausrichtung und Geschwindigkeit beobachtbar, wodurch sie im Zustandsvektor initialisiert werden. Des Weiteren wird in  $t_2$  nur seine Länge erfasst, die Breite bleibt jedoch als beobachtbar im Modell erhalten. Schließlich verlässt das Objekt in  $t_3$  den Erfassungsbereich, wodurch nur ein Punkt zur Schätzung verwendet werden kann, und die restlichen beobachtbaren Komponenten die Prädiktion einnehmen.

Die vorgestellten Beobachtbarkeitsbetrachtungen basieren auf dem Modell von Stüker [144] für den Kalman-Filter, weisen jedoch folgenden wichtigen Mehrwert auf:



**Abbildung 7-6:** Beispiel für die Objektverfolgung mit Betrachtung der variablen Referenz und Beobachtbarkeit. Zu jedem Zeitpunkt gehört die Prädiktion  $\hat{\mathbf{x}}(t_i|t_{i-1})$ , entspricht hier auch  $\hat{\mathbf{z}}(t_i)$ , Beobachtung  $\mathbf{z}(t_i)$  und die Schätzung  $\hat{\mathbf{x}}(t_i|t_i)$ . Gestrichelte Bereiche stellen unbeobachtbare Komponenten dar. Eingerahmte Komponenten sind für die Assoziation (mit Referenz  $r_{assoc}$ ) ausgewählt.

- Explizite Anpassung des Konzepts der Beobachtbarkeitsmatrix für die Beobachtungen  $\mathbf{V}_{zs}(t_i)$ .
- Bestimmung der Beobachtbarkeit der einzelnen Komponenten (direkt oder indirekt beobachtbar) des Zustandsvektors aus der Beobachtbarkeit der eingehenden Beobachtungen und umgekehrt.
- Diese Betrachtungen wurden für den Informationsfilter angepasst und ausführlich formuliert.

## 7.7 Zuverlässigkeit einer Objekthypothese

Die Zuverlässigkeit einer Objekthypothese gibt einen Maß, mit welcher Wahrscheinlichkeit das Objekt tatsächlich existiert. In anderen Arbeiten wird dies als Glaubwürdigkeit, Bedeutsamkeit, Existenzwahrscheinlichkeit oder Konfidenz der Objekte bezeichnet.

Die Zuverlässigkeit einer Objekthypothese wird hier ausschließlich aus der Erfassung des Objekts in den letzten Messzyklen geschätzt und als Zahl zwischen 0 und 1 ausgedrückt. Dafür sind vier mögliche Zustände *st* definiert und in Form eines endlichen Zustandsautomaten aufgebaut. Die Zustände sind: neue (*new*), sporadische (*sporadic*), feste (*fix*) und zu löschende (*delete*) Objekthypothese. Abbildung 7-7 stellt den Zustandsautomaten dar.

Die Zustandsübergänge sind durch die folgenden Kenngrößen bestimmt:

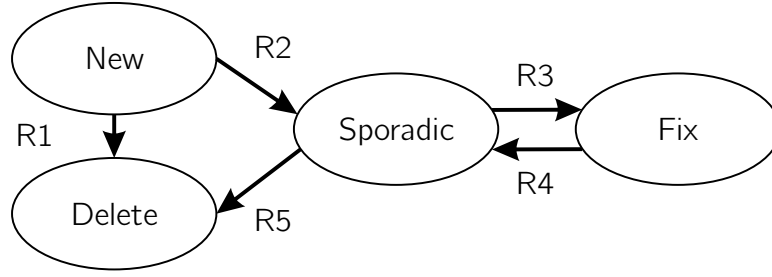
- $k_1$ : Gesamtalter der Objekthypothese;
- $k_2$ : Erfassungsfrequenz der Objekthypothese;
- $k_3$ : Zeit seit der letzten Beobachtung;

Dadurch können für jeden Zustandsübergang die Bedingungen (Regeln) anhand der Parameter  $p_{Ri}$  und  $q_{Ri}$  gegeben werden, die auf Erfahrungswerten beruhen:

**Tabelle 7-1:** Regeln zum Wechseln des Objektszustands *st*

Regel	Ist-Zustand	Bedingung	Übergang zu
R1	<i>new</i>	$k_3 > p_{R1}$	<i>delete</i>
R2	<i>new</i>	$k_3 = 0$ und $k_1 > p_{R2}$	<i>sporadic</i>
R3	<i>sporadic</i>	$k_2 > p_{R3}$ und $k_1 > q_{R3}$	<i>fix</i>
R4	<i>fix</i>	$k_3 > 0$ und $k_2 < p_{R4}$	<i>sporadic</i>
R5	<i>sporadic</i>	$k_3 > 0$ und $k_2 < p_{R5}$	<i>delete</i>

In diesem Zustandsautomaten wird eine gewisse Hysterese eingeführt. So ist normalerweise  $p_{R3}$  größer als  $p_{R4}$ . Außerdem ist normalerweise  $p_{R2}$  kleiner als  $q_{R3}$ , da ansonsten der Zustand *sporadic* übersprungen würde. Aus demselben Grund soll  $p_{R4}$  größer sein als  $p_{R5}$ . Je nach den Parametern kann eine Schwingung zwischen den Zuständen *sporadic* und *fix* entstehen, die



**Abbildung 7-7:** Zustandsautomat einer Objekthypothese.

durch entsprechende Anpassung ausgeschlossen werden muss (bspw. durch leichte Absenkung von  $p_{R3}$  oder Erhöhung von  $p_{R4}$ ).

Aus dem Zustand sowie den Kenngrößen  $k_1$ ,  $k_2$  und  $k_3$  lässt sich für jede Objekthypothese dessen momentane Zuverlässigkeit (*confidence*, *conf*) schätzen. Je höher der Zustand und die Kenngrößen  $k_1$  und  $k_2$  sind, und je niedriger die Kenngröße  $k_3$  ist, desto höher wird die Zuverlässigkeit eingestuft. Dazu kann beispielsweise folgender Ausdruck in Abhängigkeit des Zustands  $st$  verwendet werden:

$$conf = \begin{cases} 0,5 \cdot \left[ \frac{1}{3} \left( \frac{k_{1,new}}{p_{R2}} \right) + \frac{1}{3} \cdot 0 + \frac{1}{3} \left( 1 - \frac{k_{3,new}}{p_{R1}} \right) \right] & st = new, \\ 0,7 \cdot \left[ \frac{1}{3} \left( \frac{k_{1,sp}}{q_{R3}} \right) + \frac{1}{3} \left( \frac{k_{2,sp} - p_{R5}}{p_{R3} - p_{R5}} \right) + \frac{1}{3} \left( 1 - \frac{k_{3,sp}}{p_{R1}} \right) \right] & st = sporadic, \\ 1,0 \cdot \left[ \frac{1}{3} \cdot 1 + \frac{1}{3} \left( \frac{k_{2,fix} - p_{R4}}{p_{R3} - p_{R4}} \right) + \frac{1}{3} \left( 1 - \frac{k_{3,fix}}{p_{R1}} \right) \right] & st = fix, \\ 0,0 & \text{sonst.} \end{cases} \quad (7.47)$$

Dabei ist für jeden Zustand die Form gegeben:

$$conf_{st} \cdot \left[ \frac{1}{3} conf_{k_1} + \frac{1}{3} conf_{k_2} + \frac{1}{3} conf_{k_3} \right] \quad (7.48)$$

$conf_{st}$  ist eine festgelegte Grundkonfidenz des Zustands: 0,5 für neue Objekte, 0,7 für sporadische und 1,0 für feste Objekte. Diese wird mit den gemittelten Teilkonfidenzen aus den drei Kenngrößen  $k_1$ ,  $k_2$  und  $k_3$  multipliziert. Jede dieser Teilkonfidenzen ergibt sich aus der Gewichtung zwischen dem aktuellen Wert der Kenngröße und dem je nach Zustand möglichen Wertebereich. Diese kann als die Nähe zum höheren (1) bzw. niedrigeren (0) Zustand verstanden werden (vgl. Tabelle 7-1). Dazu sind die Kenngrößen wie folgt normiert, damit die Werte der Teilkonfidenzen im Bereich Null bis Eins bleiben:  $k_{1,new} = \min(k_1, p_{R2})$ ,  $k_{1,spor} = \min(k_1, q_{R3})$ ,  $k_{2,spor} = \max(p_{R5}, \min(k_2, p_{R3}))$ ,  $k_{2,fix} = \max(p_{R4}, \min(k_2, p_{R3}))$  und  $k_{3,new} = k_{3,spor} = k_{3,fix} = \min(k_3, p_{R1})$ .

## 7.8 Verwaltung und Umgang mit inkonsistenten Objekthypothesen

Ein Multi-Target-Tracking-System, wie in Abbildung 7-2 auf Seite 68 abgebildet, benötigt Verwaltungsmechanismen, die über die Zustandsschätzung der verfolgten Objekthypothesen hinausgehen. Unter anderem muss gewährleistet werden, dass neue Objekte im System initialisiert und alte Objekthypothesen aus dem System gelöscht werden. Des Weiteren sollen Inkonsistenzen in den Objekthypothesen oder zwischen Objekthypothesen und Objektbeobachtungen eingefangen und korrigierende Maßnahmen getroffen werden. Diese Inkonsistenzen zeigen sich beispielsweise bei manövrierenden Objekten, Ausreißern, und *split*- und *merge*-Effekten.

### 7.8.1 Initialisierung und Löschen von Objekthypothesen

Wurde eine Beobachtung nicht assoziiert, so wird mit dieser eine neue Objekthypothese im System initialisiert. Dazu werden die Informationsstrukturen und die Beobachtbarkeitsmatrix zu Null gesetzt. Anhand der Beobachtung wird dann die erste Schätzung durchgeführt, womit alle beobachtbaren Strukturen initialisiert werden, einschließlich der Referenz.

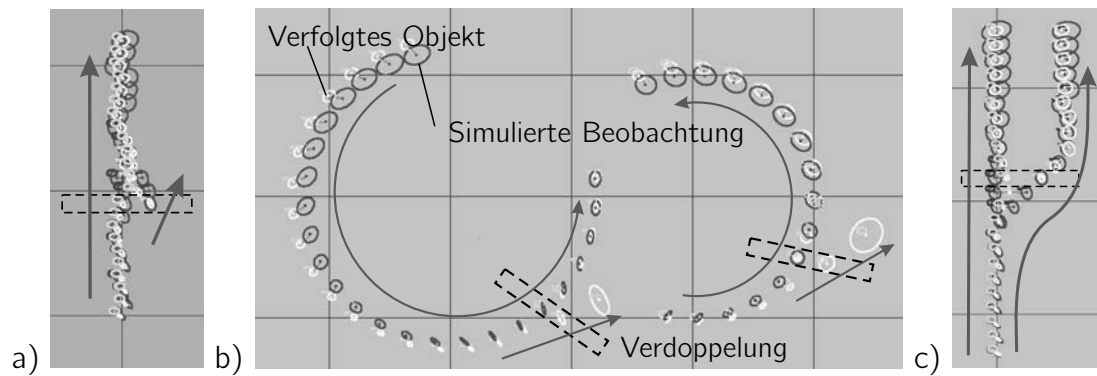
Ist die Zuverlässigkeit (siehe letzten Abschnitt) oder die Genauigkeit einer Objekthypothese sehr gering, so wird sie aus dem System gelöscht.

### 7.8.2 Verdoppelte-Hypothesen-Modellierung

Ein zentraler Aspekt dieser Arbeit liegt im Umgang mit den Objekthypothesen bei unklarer Assoziation, die auf Seite 76 eingeführt wurde. Unklare Assoziationen ergeben sich aus folgenden Gründen: wegen des großen Abstands zwischen einer Objekthypothese und einer Objektbeobachtung, wegen ungenauer Objektbeobachtungen oder Objekthypothesen, wegen Geschwindigkeits- oder Ausrichtungsänderungen der Objekte (Manövrieren), aufgrund von unterschiedlichen erfassten Ausdehnungen oder von *split*- und *merge*-Effekten.

In den erwähnten Situationen ergibt sich bei der Assoziation nicht eindeutig, ob die Objekthypothese auf Basis der Beobachtung geschätzt werden soll oder nicht. Das vorgeschlagene Verfahren besteht in der Verdoppelung von Objekthypothesen bei unklarer Assoziation, womit die Objekthypothesen  $o_1$  und  $o_2$  entstehen. Dabei wird eine der Objekthypothesen assoziiert ( $o_1$ ), während die andere unassoziiert bleibt ( $o_2$ ). Zur deutlichen Unterscheidung zwischen den Doppelobjekten wird das Prozessrauschen des assoziierten Objekts  $o_1$  künstlich erhöht (vgl. *fudge factor* von Bar-Shalom [10]), solange die Verdoppelung besteht. Damit wird erreicht, dass das assoziierte Objekt sich stark auf diese einstimmt (z.B. seine Geschwindigkeit ändert, seine Größe anpasst etc.). Beide Objekthypothesen werden als *duplicate* markiert, um rekursive Verdoppelungen zu vermeiden. Die Entscheidung, ob eins der verdoppelten Objekte beibehalten wird oder nicht, wird damit auf die nächsten Zyklen verlagert. Ist ein als *duplicate* markiertes Objekt seit  $t_d$  Millisekunden nicht beobachtet (nicht bestätigt) worden, so wird dieses auf *delete* gesetzt. Wird es jedoch beobachtet, so wird die Markierung *duplicate* entfernt und das Objekt im System beibehalten. Dabei werden die Objekt-Ids so verwaltet, dass die ursprüngliche (älteste) Id beibehalten wird. Dies ermöglicht der späteren Anwendung eine eindeutige Verfolgung des Objektes anhand seiner Id.





**Abbildung 7-8:** Beispiele für Objektverfolgung mit Verdoppelungsstrategie anhand von simulierten Beobachtungen. a) kurzzeitiger Ausreißer; b) zwei Beispiele, an denen das dynamische Modell nicht eingehalten werden konnte, was durch Verdoppelung gelöst wurde; c) Korrektes *Split* eines Objekts.

Bei der Verdoppelung können sich bei diesem Verfahren drei Situationen ergeben, die in Abbildung 7-8 anhand von simulierten Daten verdeutlicht sind:

1. Die Verdoppelung wurde durch einen Ausreißer verursacht und die zuerst assoziierte Hypothese  $o_1$  wird nach einigen Zyklen wieder gelöscht während die andere  $o_2$  beibehalten wird (Abbildung 7-8.a).
2. Das Objekt oder das eigene Fahrzeug befindet sich in einem Manöver, was das Dynamikmodell nicht erfassen kann, wodurch nur das Assoziierte  $o_1$  beibehalten wird. Die nichtassoziierte Hypothese  $o_2$  bleibt beim falschen Dynamikmodell (aus der Tangente) und wird in wenigen Zyklen gelöscht. Siehe als Beispiel dazu Abbildung 7-8.b, wo kreisförmige Bewegungen trotz des einfachen Dynamikmodells verfolgt werden können.
3. Bei der Beobachtung handelt es sich eigentlich um ein neues Objekt, welches durch Verschmelzung oder Verdeckung bisher nicht erkannt wurde. Daher werden beide beibehalten und getrennt mit jeweiligen Ids verfolgt (Abbildung 7-8.c).

Mit dieser Strategie kann das System in den meisten Situationen, die den beschriebenen Fällen entsprechen, richtig reagieren. Allerdings soll hierbei auf den hinzugekommenen Overhead hingewiesen werden, der gegebenenfalls bei sehr verrauschten Sensordaten zu einer momentan großen Anzahl von Objekthypothesen führen kann. Dies kann zu Ressourcenproblemen führen, was aufgrund immer leistungsfähiger werdender Rechner im Fahrzeug jedoch weniger problematisch ist. Darüber hinaus können falsche Verdoppelungen zu Beeinträchtigungen anderer richtiger Objekthypothesen führen, was schwerwiegender ist. Solche Situationen sollten jedenfalls durch eine frühzeitige Erkennung und Ausfilterung von sehr verrauschten Messzyklen vermieden werden.

### 7.8.3 Split und Merge von Objekten

Eines der Ziele der Fusion besteht in der Erkennung bzw. Vermeidung von fälschlicherweise zusammengefügt (*merge*) und gesplitteten Objekten (*split*). In dieser Arbeit wird diese Problematik anhand von unterschiedlichen Strategien behandelt, die im Folgenden erläutert werden.

Das *merge* von zwei realen Objekten in eine Objekthypothese kann fälschlicherweise geschehen, wenn sie sich sehr nah beieinander befinden. Dies wird zum Teil bereits durch die Assoziation und das Gating vermieden. Werden zum Beispiel zwei im Modell existierende Objekthypothesen kurzzeitig als eine einzige Beobachtung erfasst, so werden die beiden damit assoziiert (mehrfache Assoziation) und anhand der jeweiligen gemeinsamen Referenz geschätzt (Abbildung 7-9.a- $t_1$ ). Dabei wird die Dimension der beiden Objekte kurzzeitig falsch geschätzt, was nach einigen Messzyklen wieder korrigiert wird<sup>14</sup>. Kommt jedoch ein *merge* von Hypothesen vor (beide Objekte werden nur durch eins im Modell dargestellt), so wird die vorhin vorgestellte Verdoppelungsstrategie dafür verwendet, sie in weiteren Zyklen wieder zu trennen (Abbildung 7-9.b).

Das *split* einer einzelnen Objekthypothese aufgrund von mehreren Beobachtungen wird ebenso durch die mehrfache Assoziation vermieden. Wird ein einzelnes Objekt bedingt durch Sensorfehler oder Objektartefakte kurzzeitig als zwei Teilobjekte erfasst, so werden beide Beobachtungen an der jeweiligen Referenz mit der Objekthypothese assoziiert und geschätzt<sup>15</sup> (siehe Abbildung 7-9.a- $t_2$ ). Damit wird eine einzelne Objekthypothese verfolgt. Entstehen als Objektinitialisierung oder zu unterschiedlichen Beobachtungen (Verdoppelung) fälschlicherweise zwei Hypothesen für ein Objekt im Laufe mehrerer Messzyklen, so müssen diese gesondert vereinigt werden. Um zu prüfen, ob dieser Fall eintritt, wird in einem zusätzlichen Schritt die Assoziation (und Gating) innerhalb der Objekthypothesen (Objekthypothese zu Objekthypothese) durchgeführt, wie in Abschnitt 7.4. Ist die Assoziationsdistanz zwischen zwei Objekthypothesen besonders gering (nach einem sehr kleinen Schwellwert), so werden diese zusammengefügt (siehe Abbildung 7-9.c). Dazu werden sie zuerst auf eine gemeinsame Referenz gesetzt (wie in Abschnitt 7.3). Die Kovarianzmatrix und der Zustandsvektor werden dann nach der entsprechenden Genauigkeit gewichtet wie folgt berechnet<sup>16</sup>:

$$\mathbf{P}(t_i) = [\mathbf{P}_1^{-1}(t_i) + \mathbf{P}_2^{-1}(t_i)]^{-1} \quad (7.49)$$

$$\mathbf{x}(t_i) = \mathbf{P}(t_i) \cdot [\mathbf{P}_1^{-1}(t_i) \cdot \mathbf{x}_1(t_i) + \mathbf{P}_2^{-1}(t_i) \cdot \mathbf{x}_2(t_i)] \quad (7.50)$$

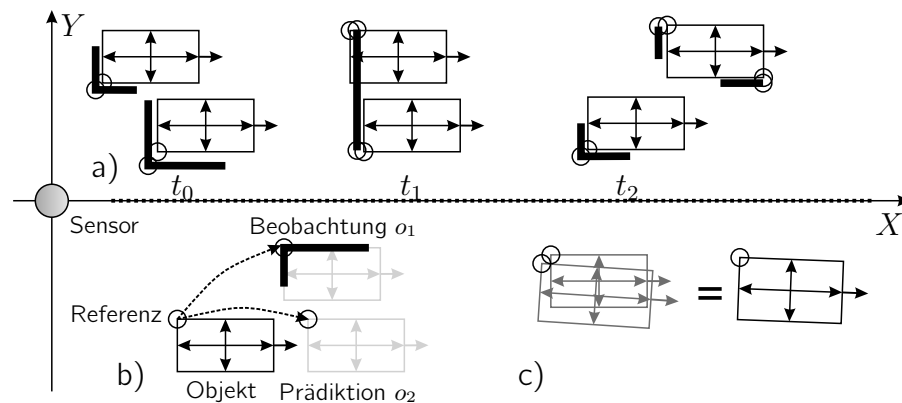
Zusätzlich muss auf die Beobachtbarkeit der zu vereinigenden Objekte geachtet werden. Deswegen werden die Gleichungen (7.49) und (7.50) in drei Schritten angewendet, und zwar für die gemeinsam beobachtbaren Komponenten und jeweils für die nur in einem Objekt beobacht-

---

<sup>14</sup>In solchen Fällen könnten die Objektdimensionen als unbeobachtbar gesetzt werden, um die falsche Schätzung zu vermeiden.

<sup>15</sup>Um die Schätzung nicht zu verfälschen, kann der Beitrag dieser beiden Beobachtungen, die ein tatsächliches Objekt darstellen, im Filter gemittelt werden.

<sup>16</sup>Analog zur Summe der Beiträge im Informationsfilter, Ausdrücke (7.23) und (7.24).



**Abbildung 7-9:** Beispiele für Vermeidung von *split* und *merge* von Objekten. a) Vermeidung von *merge* und *split* durch mehrfache Assoziation pro Beobachtung ( $t_1$ ) bzw. pro Objekthypothese ( $t_2$ ); b) Verdoppelung einer Objekthypothese, das zwei darstellte; c) Vereinigung von zwei Objekthypothesen, die eins darstellten.

baren Komponenten. Diese Strategien sind allgemein, womit jegliche (oft spezifische) Regeln vermieden werden.

Durch diese Strategien werden die Effekte von *split* und *merge* aufgehoben. Die Grenzen der Vermeidung von *splitting* zeigen sich bei voneinander entfernten Objektteilen, wie es bei vom Sensor weit entfernten Objekten vorkommen kann. Die Vermeidung von *merging* hat nur in extremen Situationen Schwierigkeiten, bei denen sich Objekte sehr nah aneinander und weit entfernt vom Sensor befinden. In der Regel sind solche Situationen sehr kurz und die Objekte weit entfernt, so dass sie weniger bedeutend bezüglich der Anwendungsebene sind.

### 7.8.4 Weitere Punkte

Es gibt eine Vielzahl an Optimierungen und Anpassungen, mit denen das Verhalten des Multi-Target-Tracking-Systems an die Sensoren und die Anwendung optimal angepasst werden kann. Es werden hier zwei weitere Punkte angesprochen: die Rücksetzung der Beobachtbarkeit und die Ungewissheit zwischen Genauigkeit und Ausdehnung der Objekte.

Die in Abschnitt 7.6 vorgestellten Strategien zur Behandlung der Beobachtbarkeit von Objekthypothesen erlauben lediglich, dass die Beobachtbarkeit erweitert wird. Im Umgang mit realen Daten zeigt sich jedoch, dass es in manchen Situationen zu falschen Schätzungen der Objektbreite bzw. Objektlänge<sup>17</sup> kommen kann. Daher wird in jedem Zyklus zusätzlich geprüft, ob die geschätzte Objektausdehnung durch die Sensoren beobachtet werden müsste und ob diese beobachtet wurde, ähnlich wie in den Beobachtbarkeitsprüfungen des Laserscanners (vgl. Abschnitt 6.2). Wird eine erwartete beobachtbare Komponente nicht beobachtet (z.B. Breite oder Länge), so spricht dies dafür, dass die Komponente eigentlich unbeobachtbar sein müsste und wird daher auf Null gesetzt (und entsprechend die Komponenten des Zustandsvektors und der Fehlerkovarianzmatrix).

Der zweite behandelte Punkt betrifft die Ungewissheit, ob durch zwei gleichzeitige Beobachtungen eines Objekts seine Position verfeinert oder seine Ausdehnung angegeben wird. Insbesondere wird dies bei der mehrfachen Assoziation relevant. Zum Lösen dieses Problems wird in

<sup>17</sup>Zum Beispiel aus dem falschen *merge* von einem Objekt mit einem Ausreißer.

dieser Arbeit als Heuristik die Beobachtung mit kürzester Distanz zur Objekthypothese (beste Assoziation) stets zur Schätzung des Objektzustands verwendet (Neuschätzung der Position und Erhöhung der Positionsgenauigkeit, wie bisher vorgestellt). Die folgenden assoziierten Beobachtungen werden zur Schätzung der Objektausdehnung verwendet. Dazu wird jede dieser Beobachtungen so verändert, dass ihre Ausdehnung die Objekthypothese mit einschließt. Die Messunsicherheit der Ausdehnung wird anhand der Positionsunsicherheit berücksichtigt. Mit der angepassten Beobachtung wird der Schätzungsschritt durchgeführt. Mit dieser Strategie ist es somit möglich ausgedehnte Objekte anhand von einzelnen Eckpunkten oder Objektteilen kooperativ zu verfolgen.

Diese Strategien haben sich in der Praxis bewährt, wie in den experimentellen Ergebnissen in Kapitel 9 gezeigt wird.

## 7.9 Zusammenfassung des Kapitels

In diesem Kapitel ist ein Fusionsansatz basierend auf der Objektebene ausführlich vorgestellt und begründet, sowie mit neuen Methoden ergänzt worden. Als Grundlage wird ein asynchroner Informationsfilter für die Fusion von Objektbeobachtungen in Objekthypothesen verwendet. Hierzu wird ein einfaches lineares Dynamikmodell ausgewählt, um die optimale Funktion des Filters zu gewährleisten. Dieses Modell ermöglicht die schnelle Filterinitialisierung und eine erfolgreiche Verfolgung in den meisten Situationen. Manövrierende Objekte können durch das Modell zum Teil nicht verfolgt werden.

Die wesentlichen neuartigen Punkte sind:

- Nutzung einer variablen Referenz der Objekte, um ausgedehnte Objekte in wechselnder Perspektive genau verfolgen zu können. Dabei wird die Größe der Objekte berücksichtigt und trotzdem ein punktförmiges Objektmodell verwendet.
- Mehrfachassoziationsverfahren anhand der asymmetrischen Mahalanobis-Distanz und Gating-Strategien, um Robustheit gegenüber Ausreißern, manövrierenden Objekten und *split*- und *merge*-Effekten zu erreichen.
- Die Filterstrukturen sind um eine effektive Verdoppelungsstrategie zur besseren Verfolgung von manövrierenden Objekten ergänzt worden. Im Gegensatz zu anderen adaptiven Filterverfahren, können hiermit Fehler in der Assoziation später behoben werden. Außerdem muss der Filter nicht auf bestimmte Manöver zugeschnitten werden.
- Vermeidung von *split*- und *merge*-Effekten mittels einer Kombination der Mehrfachassoziation und der Verdoppelungsstrategie bzw. einer Vereinigungsstrategie.
- Integration des Konzeptes der partiellen Beobachtbarkeit in Form einer Beobachtbarkeitsmatrix in Kombination mit der logischen Referenz in den Ausdrücken des Informationsfilters. Damit erhält das Fusionsystem neben dem konkurrierenden einen komplementären Anteil und Teilverdeckungen werden auch zwischen Objekten behandelt.

Mit diesem Verfahren wird eine stabile und robuste Fusion erreicht. Der Ansatz bleibt dabei flexibel und erweiterbar. Ein ausführliches und quantitatives Testen erfolgt in Kapitel 9.

## 8 Gefahrenerkennung

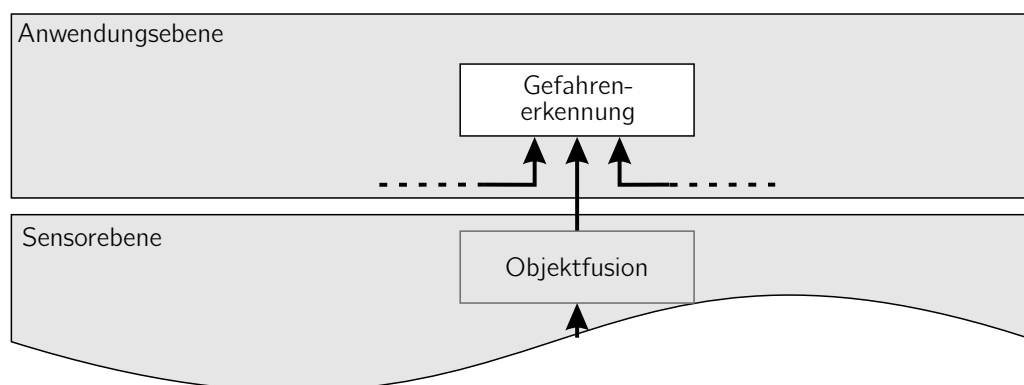
Die erkannten Informationen aus der Fahrsituation, in diesem Fall die Objektdaten, werden in der Verarbeitungskette auf der Anwendungsebene verwendet. In dieser Arbeit wird die Erkennung von potentiellen Gefahrensituationen als Anwendung betrachtet. Es wird eine technische Lösung vorgestellt, die ein Maß zur Schätzung von Gefahren liefert. Dabei wird eine Abstraktion von den verwendeten Daten erzielt, so dass der Ansatz allgemein und erweiterbar ist. Als Abstraktion wird die statistische Häufigkeit der Ereignisse betrachtet, welche atypische Situationen umfassen, da diese potentielle Gefahren darstellen können. Diese Konzepte werden anhand von Objektdaten umgesetzt, indem Indikatoren wie TTC – *Time To Collision*, Zeit bis zum Zusammenstoß – und THW – *Time Headway*, Zeitlücke – abgeleitet werden.

Abbildung 8-1 stellt den Zusammenhang der Erkennung von Gefahrensituationen mit der restlichen Arbeit dar (aus Abbildung 1-3 auf Seite 7).

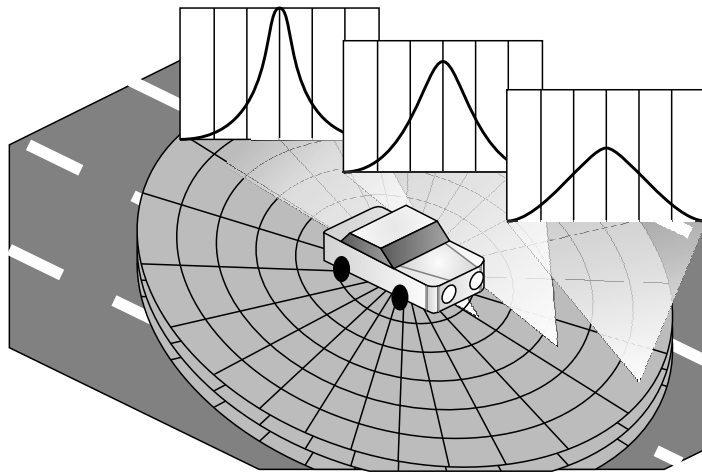
### 8.1 Statistische Gefahrenerkennung

In diesem Kapitel wird ein Ansatz vorgestellt, mit dem potentielle Gefahrensituationen in Form von atypischen Ereignissen erkannt werden können. Dafür müssen typische Situationen charakterisiert werden. Hierzu werden die *a posteriori* Wahrscheinlichkeiten der betrachteten Komponenten in Form von Häufigkeitsverteilungen aus Trainingsdaten ermittelt. Diese werden in einer Datenstruktur gespeichert, dem statistischen Informationsgitter (SIG). Diese Struktur ist schematisch in Abbildung 8-2 dargestellt und wird in den folgenden Absätzen erläutert.

Die komplexe Fahrsituation wird in Teilaspekte zerlegt, wie beispielsweise andere Verkehrsteilnehmer, Hindernisse, Straße, Witterung etc. Diese Teilaspekte werden in unterschiedlichen Schichten des SIG dargestellt und behandelt – siehe graue Scheiben in Abbildung 8-2. In der Regel sind die Teilaspekte mehr oder weniger miteinander verwoben. Zum Beispiel passen sich die Geschwindigkeiten der Witterung an oder die Bewegung der Objekte richtet sich anhand der Straßengeometrie aus. Daher müssten zu den Teilwahrscheinlichkeiten der einzelnen Aspekte stets die bedingten Wahrscheinlichkeiten zwischen ihnen betrachtet werden (bspw. die Wahrscheinlichkeit, dass der Vordermann bremst, bedingt durch die Annäherung an eine Kreuzung).



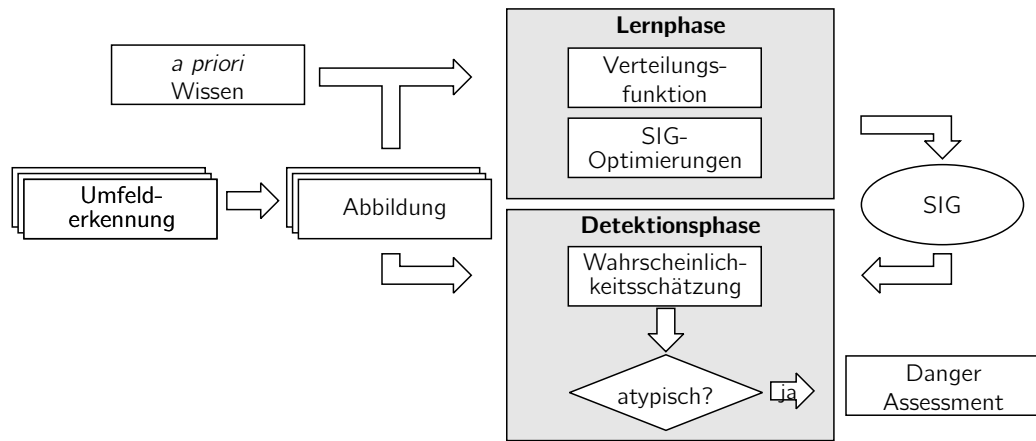
**Abbildung 8-1:** Anwendungsebene im Rahmen der Arbeit.



**Abbildung 8-2:** Graphische Darstellung des statistischen Informationsgitters (SIG).

Diese Berücksichtigung der bedingten Wahrscheinlichkeiten würde jedoch die Dimension des Ansatzes sprengen und damit für die gleiche Menge an Trainingsdaten deutlich an Signifikanz verlieren. Daher wird in dem Ansatz jeder Teilaspekt nur im Einzelnen betrachtet, wodurch jedoch eine Reduktion der Komplexität der erkennbaren Gefahren in Kauf genommen wird. Das gleichzeitige Auftreten eines atypischen Ereignisses in zwei Teilaspekten bedeutet damit nicht unbedingt, dass die Situation umso atypischer ist, denn beide könnten stark korreliert sein. Zum Beispiel stellen eine schnelle Annäherung an den Vordermann und die überhöhte Eigengeschwindigkeit möglicherweise zwei Aspekte derselben Situation dar. Durch diese Einschränkung ist die Erweiterbarkeit des Systems erleichtert, da dabei die Abhängigkeiten zwischen Teilaspekten nicht betrachtet werden. In dieser Arbeit werden als Teilaspekte die TTC und die THW beispielhaft berücksichtigt.

Wie aus dem Stand der Technik (Kapitel 2) ersichtlich wurde, besteht eine der Herausforderungen der Gefahrenerkennung in der Feststellung eines geeigneten Schwellwertes, um eine Warnung auszulösen. Idealerweise sollte dieser dynamisch, adaptiv und situationsspezifisch errechnet werden. Um dies zu erreichen, werden Fahrsituationen in dem hier vorgestellten Ansatz nach unterschiedlichen Kriterien differenziert. Der Grad an Gefahr eines bestimmten Ereignisses hängt somit grundsätzlich von der Straßenart ( $rt$ ), dem aktuellen Manöver ( $mn$ ) und der relativen Position der Objekte ( $xy$ ) zum eigenen Fahrzeug ab. Daher enthält das SIG  $n_{rt} \times n_{mn} \times n_{xy}$  Häufigkeitsverteilungen für jede Komponente, was eine große Datenmenge nach sich zieht. Abbildung 8-2 verdeutlicht die Ortskomponente anhand des Gitters und der unterschiedlichen dargestellten Verteilungen. Insbesondere diese Ortsabhängigkeit der Gefahrenanalyse ermöglicht in Vergleich zu herkömmlichen Ansätzen eine differenzierte und frühzeitige Feststellung von Gefahren. Daher wird der Fokus der Arbeit im Folgenden darauf gesetzt. Eine Gefahr, die sich schon in einer großen Entfernung entwickelt, kann durch einen angepassten (höheren) Schwellwert der TTC früher festgestellt werden, als mit einem festen Schwellwert. Damit die Ortsbetrachtung allgemeingültig sein kann, sollten nicht kartesische oder polare Koordinaten sondern straßenbezogene Koordinaten mit Zentrum im Ego-Fahrzeug verwendet werden. Dabei verläuft die  $X$ -Koordinate stets entlang der Straße (unabhängig von Kurven) und die  $Y$ -Koordinate in der Straßenquerrichtung, vgl. Eidehall [47] (*Road-Aligned Coordinates, RAC*) und Wohlfeil [166]. Wird der Ansatz auf Landstraßen oder innenstädtische Situationen angewendet, so wäre eine weitere Koordinate für Querstraßen angebracht.



**Abbildung 8-3:** Prozessüberblick für die Gefahrenerkennung.

Es gibt in statistisch motivierten Ansätzen in der Regel zwei Phasen, welche in diesem Ansatz auch zu finden sind: das Anlernen sowie Charakterisieren von typischen Ereignissen und die Detektionsphase. Beide sind in Abbildung 8-3 dargestellt. In der Lernphase wird mithilfe von vielen Testdaten oder *a priori* Wissen die Wahrscheinlichkeitsverteilung<sup>1</sup> der betrachteten Komponenten ermittelt und im SIG gespeichert. Dafür ist ein Schritt zur Abbildung der Informationen der Fahrsituation in das SIG notwendig (bspw. Berechnung der TTC aus Objektdaten). Zusätzlich können in der Lernphase die SIG-Strukturen optimiert werden, wie beispielsweise durch Glättung oder Datenkompression. In der Detektionsphase werden die Daten der Fahrsituation (als Ereignis  $z_i$  betrachtet) in das SIG abgebildet und deren Wahrscheinlichkeit  $P(z_i)$  mithilfe der trainierten Verteilungen im SIG geschätzt.

Ausgehend von der Definition einer Zelle  $c$  des SIG<sup>2</sup> als Häufigkeitshistogramm

$$SIG[c, z_i] \triangleq \text{Anzahl Trainingswerte für Ereignis } z \text{ in Zelle } c, \quad (8.1)$$

wobei das Ereignis einem Intervall  $z_i \in (z_{inf}, z_{sup}]$  des Histogramms (Spalte) entspricht, wird die Wahrscheinlichkeit  $P(z_i)$  eines Ereignisses  $z_i$  in der Zelle  $c$  wie folgt berechnet

$$P(c, z_i) = \frac{SIG[c, z_i]}{\sum_{i'} SIG[c, z_{i'}]} \quad (8.2)$$

Besitzt diese Wahrscheinlichkeit einen sehr kleinen Wert, so wird die Situation als atypisch eingestuft und muss auf eine potentielle Gefahr hin geprüft werden (*danger assessment*). Der dazu verwendete Schwellwert bezieht sich auf die Wahrscheinlichkeit und ist somit von den zugrundeliegenden Daten abgekoppelt. Er kann daher fest sein und besitzt trotzdem ein adaptives Verhalten<sup>3</sup>. Zum Beispiel ergibt ein Schwellwert von 0,01 eine Warnung bei einer niedrigeren

<sup>1</sup>Hier als Häufigkeitsverteilung gespeichert.

<sup>2</sup>Die Zelle ergibt sich aus den vorhin erwähnten Kriterien *rt*, *mn* und *xy*.

<sup>3</sup>Vergleiche andere Ansätze zur Festlegung des Schwellwertes, wie Eidehall [47] oder Yang et al. [168, 169],

TTC in 20 m als in 70 m Entfernung, da bei größeren Entfernungen auch höhere TTC-Werte normal sind.

Bei einer ausreichend hohen Granularität der Häufigkeitshistogramme tendiert die Wahrscheinlichkeit  $P(c, z_i)$  jedes Ereignisses  $z_i$  – im Folgenden vereinfachend als  $P(z_i)$  ausgedrückt – zu Null. Im vorgestellten Ansatz werden die zu betrachtenden Komponenten so ausgewählt, dass deren Wahrscheinlichkeitsdichtefunktionen ein einziges Maximum besitzen, so dass atypische Ereignisse nur am Funktionsrand stattfinden können. Dank dieser Einschränkung können statt der Wahrscheinlichkeiten einzelner Ereignisse die Randwahrscheinlichkeiten als Kriterium für Seltenheit verwendet werden. Statt der Häufigkeitsfunktion (äquivalent zur Dichtefunktion) wird die (kumulative) Verteilungsfunktion  $F(c, z_i)$  – oder vereinfacht  $F(z_i)$  – betrachtet, die für eine SIG-Zelle  $c$  folgendermaßen definiert ist:

$$F(z_i) = P(z \leq z_i) = \frac{\sum_{i'=0}^i \text{SIG}[c, z_{i'}]}{\sum_{i'} \text{SIG}[c, z_{i'}]} \quad (8.3)$$

Damit sind die Schwellwerte  $\alpha$  durch die Quantile  $F^{-1}(\alpha)$  und  $F^{-1}(1 - \alpha)$  gegeben, definiert als:

$$F^{-1}(\alpha) \triangleq z_i : F(z_i) > \alpha \text{ und } F(z_{i-1}) \leq \alpha \quad (8.4)$$

Die Komponenten des SIG werden so ausgewählt, dass sie eine direkte Beziehung zur Gefahr haben (wie TTC, THW etc., nicht jedoch die eigene Geschwindigkeit oder Position). Damit stellt ein atypisches Ereignis stets eine Gefährdung dar und die berechnete Randwahrscheinlichkeit  $F(z_i)$  entspricht dabei der Sicherheit der Erkennung. Der Schritt des *danger assessment* ist somit im Rahmen dieser Arbeit nicht mehr notwendig. Eine Plausibilisierung mit weiteren Komponenten sowie ein Vergleich mit der Fahrerwahrnehmung werden hier nicht weiter ausgeführt.

## 8.2 Praktische Umsetzung der Gefahrenerkennung

Für die prototypische Implementierung eines SIG wird ein praktisches Szenario vorgestellt, das auf den Objektdaten aus den in den letzten Kapiteln vorgestellten Erkennungs- und Verfolgungsmethoden basiert. Dabei wird der Fokus zunächst auf Autobahnszenarien gelegt.

Die Objektdaten werden in zwei unterschiedliche SIG-Schichten abgebildet, nämlich als TTC und als THW. Diese werden in der Literatur häufig als Indikatoren für Gefährdungen verwendet. Vogel hat in [153] die beiden Indikatoren verglichen und ist zu dem Schluss gekommen, dass beide unabhängig voneinander sind. Dabei stellt eine kleine THW eine potentielle Gefährdung und eine kleine TTC eine kritische Gefährdung dar. Im SIG ist eine ausführliche Auswertung dieser

---

diskutiert in Abschnitt 2.5.



Indikatoren nach relativer Position zu anderen Objekten bzw. Fahrzeugen möglich. Die TTC gibt Aufschluss über typische Relativgeschwindigkeiten zu einem Vordermann in Abhängigkeit von dessen Entfernung. Die THW bezieht sich auf die eigene Geschwindigkeit in Abhängigkeit von der Position des Vordermanns.

Für die TTC liegt der Fokus sowohl auf dem eigenen Fahrstreifen als auch auf benachbarten Fahrstreifen oder Hindernissen an den Straßenseiten. Daher soll das SIG in der Querrichtung in drei Segmente eingeteilt werden. In Längsrichtung werden die Informationen in feste Abstände zerlegt (bspw. 10 m). Die THW, als abgeleitete Größe des Folgeverhaltens, ist nur für den eigenen Fahrstreifen sinnvoll. Daher werden die Informationen ausschließlich in der Längsrichtung betrachtet und in feste Abstände zerlegt. Zur Unterteilung der Häufigkeitsverteilung werden hier Schritte mit Größe 0,5 s verwendet,  $z_i = (0,5 \cdot i, 0,5 \cdot i + 0,5]$ ,  $i \in \mathbb{N}$ .

Im Folgenden wird auf die Schritte eingegangen, die zur praktischen Umsetzung des SIG notwendig sind. Dies beinhaltet eine Koordinatentransformation, die Abbildung der Objektdaten in das SIG (hier als TTC und THW), die Anwendung der Daten zum Trainieren der SIG-Struktur und die Verwendung der trainierten Daten zur Erkennung von atypischen Ereignissen. In Kapitel 9 werden einige Details zur prototypischen Implementierung und zum Testen von Teilen dieser Methode anhand von Simulationsdaten vorgestellt.

### 8.2.1 Umrechnung in straßengeführte Koordinaten

Die erkannten Objektdaten liegen in fahrzeugzentrierten lokalen Koordinaten vor. Mit Kenntnis der Straßengeometrie und der eigenen Position sollen diese in straßengeführte Koordinaten (*Road-Aligned Coordinates*, vgl. Eidehall [47]) umgerechnet werden. Hierzu können beispielsweise die von Eidehall [47] vorgestellten Ausdrücke verwendet werden. Zuerst werden die lokalen Objektdaten  $(x_{veh,l}, y_{veh,l}, \dot{x}_{veh,l}, \dot{y}_{veh,l})$  anhand der Position und Geschwindigkeit des eigenen Fahrzeugs  $(x_{ego,w}, y_{ego,w}, \dot{x}_{ego,w}, \dot{y}_{ego,w})$  in globalen Koordinaten  $(x_{veh,w}, y_{veh,w}, \dot{x}_{veh,w}, \dot{y}_{veh,w})$  ausgedrückt (z.B. in UTM<sup>4</sup>). Die eigene Fahrzeugposition und die Objektposition werden dann anhand der Kenntnis der Straßengeometrie in straßengeführten Koordinaten,  $(x_{ego,RAC}, y_{ego,RAC}, \dot{x}_{ego,RAC}, \dot{y}_{ego,RAC})$  und  $(x_{veh,RAC}, y_{veh,RAC}, \dot{x}_{veh,RAC}, \dot{y}_{veh,RAC})$ , umgerechnet. Aus der resultierenden Koordinate in Querrichtung ( $y_{RAC}$ ) und der Kenntnis der Anzahl von Fahrstreifen und deren Breite lässt sich der Fahrstreifen jedes Objekts bestimmen.

### 8.2.2 Berechnung der TTC

Die TTC (*Time To Collision*) ist die Zeit, nach der eine Kollision mit einem anderen Fahrzeug bzw. Objekt, ausgehend von den verwendeten dynamischen Modellen, stattfinden würde, insofern sich keine Stör- oder Kontrollgrößen ändern. Die Berechnung basiert auf dem Dynamikmodell der Objekte, hier einer geradlinigen Bewegung mit konstanter Geschwindigkeit. Die TTC wird wie folgt berechnet

$$t_{TTC} = \frac{d}{v_{rel}}, \quad (8.5)$$

<sup>4</sup>UTM-Koordinatensystem (*Universal Transverse Mercator*)

wobei  $d$  die Entfernung zwischen den Objekten und  $v_{rel}$  deren Relativgeschwindigkeit ist<sup>5</sup>. Die TTC kann sowohl für den Vordermann und Hintermann als auch für Fahrzeuge in anderen Fahrstreifen berechnet werden. Dafür ist es zusätzlich notwendig zu berechnen, ob sich deren Trajektorien schneiden werden, was hier nicht weiter ausgeführt wird.

### 8.2.3 Berechnung der THW

Die THW (*Time Headway*) entspricht der Zeitlücke zwischen zwei hintereinander fahrenden Fahrzeugen. Sie ist als die Zeit definiert, die zwischen dem Erreichen eines Punktes durch das erste und das zweite Fahrzeug vergeht. Die THW wird wie folgt berechnet:

$$t_{THW} = \frac{(x_{veh,RAC} - x_{ego,RAC})}{\dot{x}_{ego,RAC}} \quad (8.6)$$

Diese Berechnung ist sowohl in Bezug zum Vordermann als auch zum Hintermann durchführbar.

### 8.2.4 Erzeugung der SIG-Struktur

In der Lernphase werden die Daten für jeden Messzyklus in der entsprechenden SIG-Struktur aufsummiert. Dies soll auf der Grundlage der Datengenauigkeiten passieren, die in Form einer Fehlerkovarianzmatrix  $\mathbf{P}$  gegeben ist. Diese Fehlerkovarianz soll zunächst von lokalen Objektkoordinaten auf straßengeführte lokale Koordinaten transformiert werden. Dafür wird die Varianzfortpflanzung auf die Transformationen aus Abschnitt 8.2.1 angewendet. Da es sich um nichtlineare Funktionen handelt, werden deren Jacobi-Matrizen als Design-Matrizen verwendet. Die erhaltene Kovarianzmatrix wird wiederum mit dem Ausdruck (8.5) für die TTC transformiert, hier als Taylor-Reihe ausgedrückt:

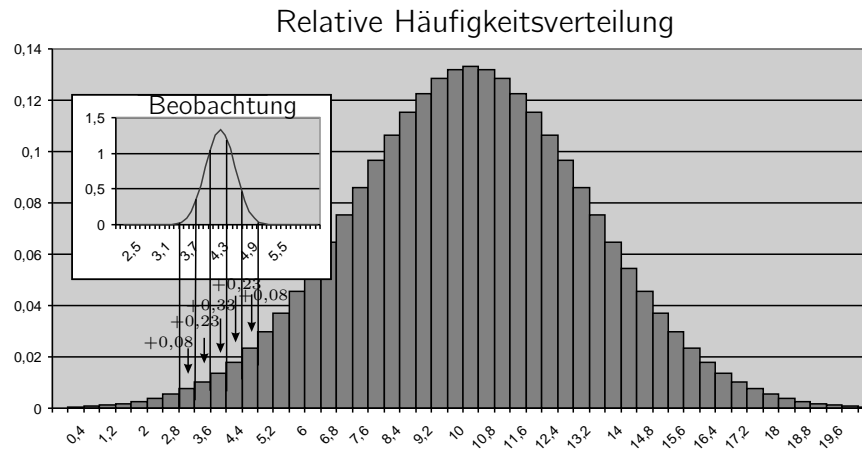
$$[t_{TTC}] = t_{TTC,0} + \mathbf{A}_{TTC1} \begin{bmatrix} x_{veh,RAC} \\ y_{veh,RAC} \\ \dot{x}_{veh,RAC} \\ \dot{y}_{veh,RAC} \end{bmatrix} + \mathbf{A}_{TTC2} \begin{bmatrix} x_{ego,RAC} \\ y_{ego,RAC} \\ \dot{x}_{ego,RAC} \\ \dot{y}_{ego,RAC} \end{bmatrix} \quad (8.7)$$

$$\mathbf{P}_{TTC} = \mathbf{A}_{TTC1} \cdot \mathbf{P}_{veh,r} \cdot \mathbf{A}_{TTC1}^T + \mathbf{A}_{TTC2} \cdot \mathbf{P}_{ego,r} \cdot \mathbf{A}_{TTC2}^T \quad (8.8)$$

Für die THW wird analog mit dem Ausdruck (8.6) verfahren.

Aus diesen Berechnungen erhält man zu jedem TTC-Wert eine Information zur jeweiligen Fehlervarianz,  $\sigma_{TTC}^2$ . Um diese ins SIG hinzuzufügen werden sie als Verteilung betrachtet (mit Erwartungswert  $t_{TTC}$  und Varianz  $\sigma_{TTC}^2$ ). Dabei kann diese Beobachtung möglicherweise Einfluss auf mehrere Säule der Häufigkeitsverteilung haben, wie in Abbildung 8-4 dargestellt wird. Der Effekt von qualitativ besseren Daten wird in Abschnitt 8.3 diskutiert. Diese Überlegungen werden

<sup>5</sup>Unter Betrachtung der unterschiedlichen Kurswinkel.



**Abbildung 8-4:** Hinzufügung neuer Daten in das SIG gewichtet mit deren Fehlerkovarianz.

analog für die THW angewendet.

## 8.2.5 Erkennung von atypischen Ereignissen

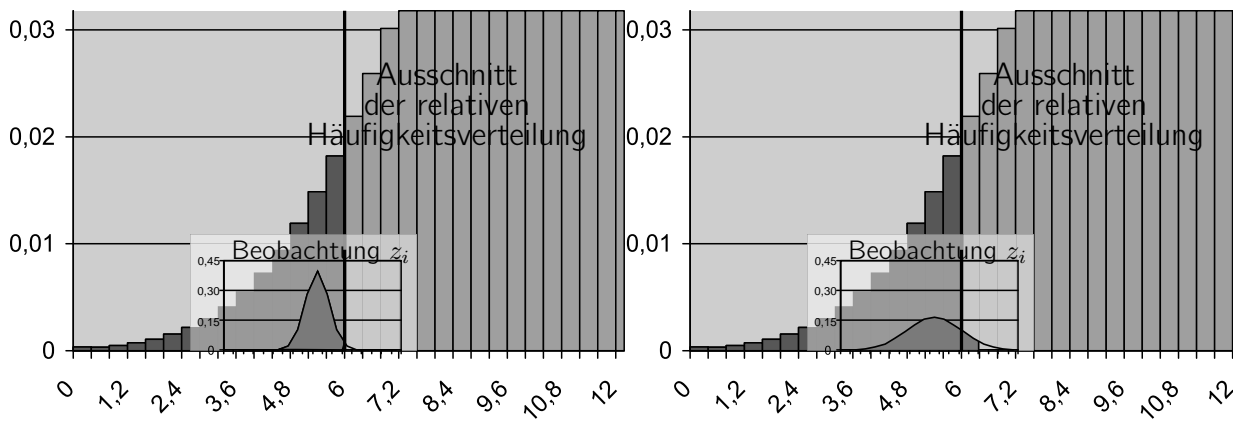
In der Detektionsphase werden die eingehenden Daten mit der im SIG bestimmten *a posteriori* Wahrscheinlichkeitsverteilung verglichen, um eine Aussage über deren Wahrscheinlichkeit zu treffen. Im idealen Fall, in dem keine Messunsicherheiten vorhanden sind, können die eingehenden Daten  $z_i$  direkt mit dem im Ausdruck (8.4) auf Seite 96 ermittelten Schwellwert  $F^{-1}(\alpha)$  verglichen werden. Da die Daten, wie im letzten Unterabschnitt berechnet, eine Messunsicherheit (Kovarianz  $\sigma_{z_i}$ ) aufweisen, muss der Vergleich auf statistischer Basis geschehen. Die zu bestätigende linksseitige Nullhypothese besagt, dass der Beobachtungswert nicht atypisch ist, also, dass der eingehende Wert  $z_i$  größer ist als der Schwellwert  $F^{-1}(\alpha)$ :

$$H_0 : z_i \geq F^{-1}(\alpha)$$

Für den Gauß-Test wird die Testprüfgröße  $z'_i$  berechnet:

$$z'_i = \frac{z_i - F^{-1}(\alpha)}{\sigma_{z_i}} \quad (8.9)$$

Diese Hypothese kann mit einem Signifikanzniveau von  $\alpha'$  abgelehnt werden (das Ereignis ist doch atypisch), wenn  $z'_i < u(\alpha')$ , wobei  $u(\alpha')$  das  $\alpha'$ -Quantil der Standardnormalverteilung  $\mathcal{N}(0, 1)$  ist. Abbildung 8-5 verdeutlicht diesen Test. In beiden Beispielen ist die Oberfläche unter der Beobachtungskurve mit Wahrscheinlichkeit  $\alpha'$  (bspw. 0,9) markiert. Im linken Beispiel liegt diese Fläche unter dem Schwellwert, womit das Ereignis eindeutig als atypisch angenommen wird. Die Beobachtung im rechten Beispiel ist wegen der größeren Messunsicherheit nicht signifikant unter dem Schwellwert. Eine höhere Messunsicherheit hat somit in der Detektionsphase zur Folge, dass die Warnung erst später erfolgt.



**Abbildung 8-5:** Statistischer Test zum Vergleich der Beobachtungen mit dem Schwellwert  $F^{-1}(\alpha)$  (schwarzer Balken) im SIG. Die linke Beobachtung würde mit hoher Wahrscheinlichkeit unter dem Schwellwert liegen und somit eine Warnung auslösen. Die rechte Beobachtung ist aufgrund ihrer Streuung nicht signifikant unter dem Schwellwert und würde daher keine Warnung auslösen.

### 8.3 Auslegung der Gefahrenerkennung und Bezug zur Güte der Objektdaten

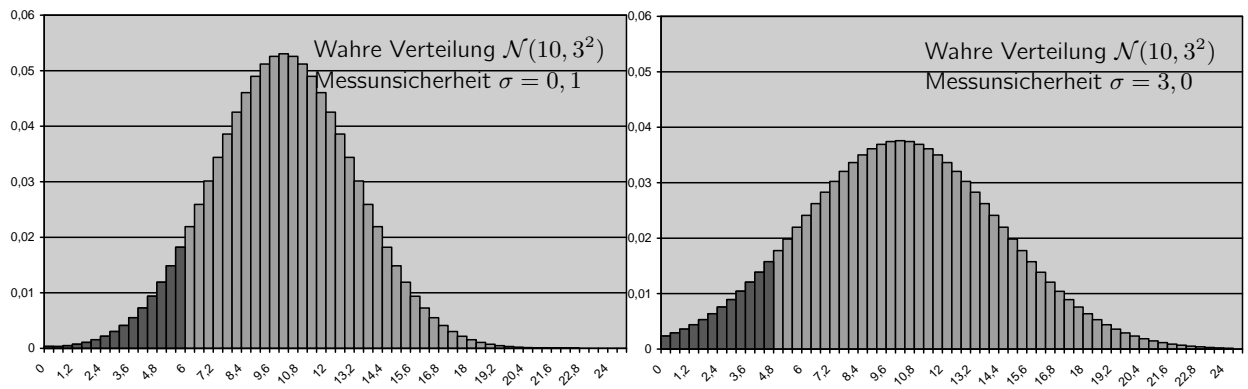
Aus dem letzten Abschnitt ist ersichtlich, dass die Qualität der Daten einen wichtigen Einfluss auf die Gefahrenerkennung hat. In diesem Abschnitt wird nun der Einfluss von Erfassungsbereich, Genauigkeit, Verfügbarkeit und Zuverlässigkeit der Objektdaten auf die Gefahrenerkennung diskutiert. Diese Aspekte werden im Folgenden getrennt und systematisch behandelt.

#### 8.3.1 Einfluss vom Erfassungsbereich der Objekterkennung auf die Gefahrenerkennung

Der Erfassungsbereich der Objekterkennung hat einen direkten Einfluss auf den Erfassungsbereich der Gefahrenerkennung. Je größer der Erfassungswinkel der Objekterkennung ist, desto mehr Deckung wird die Gefahrenerkennung haben, wie zum Beispiel im toten Winkel. Je größer die Reichweite der Erkennung ist, desto früher können atypische Situationen, bspw. von weiter entfernten Objekten, festgestellt werden. Allerdings ist der Grad an Gefahr bei atypischen Bewegungen von sehr entfernten Objekten geringer als bei demselben Schwellwert mit näheren Objekten. Das liegt daran, dass die Objekte bzw. das eigene Fahrzeug mehr Zeit zum Reagieren haben. Da in der Regel die Datengenauigkeit mit der Entfernung der Objekte absinkt, sind die Schwellwerte für den entfernten Bereich weniger empfindlich, womit Warnungen tendenziell zu spät ausgelöst werden (siehe nächsten Unterabschnitt).

#### 8.3.2 Einfluss der Objektgenauigkeit auf die Gefahrenerkennung

Eine hohe Genauigkeit der Daten ermöglicht eine bessere Gefahrenerkennung. Die Effekte einer höheren oder niedrigeren Genauigkeit lassen sich anhand zweier Beispielsimulationen zeigen, deren Ergebnisse in Abbildung 8-6 dargestellt sind. Hierbei handelt es sich um die relative Häufigkeitsverteilung für dieselbe Zelle mit jeweils unterschiedlichen Messunsicherheiten der eingehenden Beobachtungen. Die wahre simulierte Zufallsgröße hat die Verteilung  $\sim \mathcal{N}(10, 3^2)$ . Auf



**Abbildung 8-6:** Einfluss der Datengenauigkeit auf die Gefahrenerkennung. Beide Verteilungen ergeben sich aus derselben wahren Verteilung durch unterschiedliche Messunsicherheiten. Der Schwellwert für  $\alpha = 0,1$  beträgt links 6 s und rechts 4,6 s.

der linken Seite wurden die Eingangsdaten mit einer Messunsicherheit bzw. Standardabweichung von  $\sigma = 0,1$  und auf der rechten Seite mit  $\sigma = 3,0$  versehen. In der Trainingsphase wird damit die Häufigkeitsverteilung ermittelt, wie in Abbildung 8-4 gezeigt wird.

Die Genauigkeit der Daten spiegelt sich in der SIG-Verteilung wider. Je höher die Unsicherheit der Eingangsdaten ist, desto breiter ist die resultierende Verteilung. Eine breitere Häufigkeitsverteilung im SIG hat zufolge, dass sich das Quantil für einen Schwellwert  $\alpha$  (bspw. 0,1) nach links verschiebt, also dass eine Warnung erst bei extremeren Situationen (sprich später) gegeben wird, als bei genaueren Daten. Im linken Beispiel (mit höherer Genauigkeit) entspricht  $F^{-1}(\alpha)$  einer TTC von 6 s und im rechten Beispiel einer TTC von 4,6 s.

In ähnlicher Art spiegelt sich diese Problematik in der Detektionsphase wider. Sind die Daten ungenau, so ist der Vergleich mit dem Schwellwert verschwommen. Das führt dazu, dass die Warnungen (womöglich fälschlicherweise) erst bei niedrigeren Werten ausgegeben werden. Das heißt, es wird später gewarnt (siehe Überlegungen in Unterabschnitt 8.2.5).

### 8.3.3 Einfluss der Verfügbarkeit der Objektdaten auf die Gefahrenerkennung

Die Verfügbarkeit der Daten hat keinen direkten Einfluss auf die Wahrscheinlichkeitsverteilungen im SIG. Die Daten werden in der Trainingsphase mit einer niedrigeren Frequenz (bspw. 2 Hz) als die Datenverfügbarkeit (10-30 Hz) dem SIG hinzugefügt.

In der Detektionsphase weist die geringe Datenverfügbarkeit ebenso die Effekte einer geringen Datengenauigkeit auf. Außerdem addiert sich zu dem durch die Ungenauigkeit verzögerten Warnungszeitpunkt noch die längere Zykluszeit. Diese Verzögerung ist in der Regel in der Größenordnung von unter 100 ms.

### 8.3.4 Einfluss der Zuverlässigkeit der Objektdaten auf die Gefahrenerkennung

Sowohl für die Trainingsphase als auch für die Detektionsphase sollten nur Objektdaten verwendet werden, die eine hohe Zuverlässigkeit besitzen. Andernfalls besteht die Gefahr, dass

Fehlwarnungen ausgegeben oder Gefahrensituationen verpasst werden.

Die Fehldetektionsrate (*false negatives*) des Erkennungssystems hat einen direkten Einfluss auf die Gefahrenerkennung. Der Umgang mit unzuverlässigen Daten und deren Effekten können zweier Art sein. Einerseits können wichtige Objekte verpasst werden, wenn nur eine hohe Zuverlässigkeit zugelassen wird. Damit gelangen auch weniger Daten in das SIG, womit die Signifikanz der Aussagen in der Detektionsphase geringer sein wird. Andererseits kann das System offener gegenüber unzuverlässigen Daten gestaltet werden, womit *false positives* fälschlicherweise in die SIG-Strukturen gelangen könnten. In diesem Fall hängen die Effekte von der Art der *false positives* ab. Weisen diese eine normale Bewegung auf, so wird die Verteilung enger und der Schwellwert höher, womit es in der Detektionsphase zu früheren Warnungen (*false alarms*) kommen wird. Weisen die *false positives* eine atypische Bewegung auf, so werden sie trotzdem in die SIG-Struktur als typisch aufgenommen (da Trainingsphase), womit die Verteilungsform breiter und der Schwellwert niedriger wird (ähnlich wie in der rechten Seite von Abbildung 8-6). Dies wird in der Detektionsphase zu verzögerten Warnungen bei echten Gefahren führen.

## 8.4 Zusammenfassung und Diskussion der Gefahrenerkennung

In diesem Kapitel ist ein technischer Ansatz zur statistischen Erkennung von Gefahren in Fahrsituationen vorgestellt worden. Dabei werden seltene oder atypische Ereignisse als Anzeichen für eine potentielle Gefahr näher betrachtet. Die grundlegenden Strukturen sind die Häufigkeitsverteilungen, die in einem statistischen Informationsgitter (SIG) verwaltet werden. Im SIG werden die Daten nach unterschiedlichen Kriterien (z.B. Ortsabhängigkeit) aufgeschlüsselt, womit die Gefahrenerkennung selektiv gestaltet wird. So werden hier als Beispiele aus den Objektdaten die TTC und die THW in Abhängigkeit von der Objektentfernung analysiert, was einen situationsspezifischen Schwellwert ergibt. Die experimentellen Ergebnisse dazu werden in Kapitel 9 vorgestellt. Dabei liegt der Fokus in der Plausibilität des Ansatzes. Ein Vergleich mit der Fahrerwahrnehmung über die Gefahren wird im Rahmen dieser Arbeit nicht betrachtet.

Die statistische Gefahrenerkennung besteht aus zwei Phasen. In der Ersten werden Trainingsdaten verwendet, um die *a posteriori* Wahrscheinlichkeitsverteilungen im SIG zu bestimmen. In der zweiten Phase werden die aktuellen Beobachtungen mit den Verteilungen im SIG verglichen, um eine Aussage über deren Wahrscheinlichkeit zu treffen. In beiden Phasen wird die Messunsicherheit der Daten mitberücksichtigt.

Zudem wurde in diesem Kapitel über die Auslegung des SIG und über die Effekte der Datengüte auf die Gefahrenerkennung diskutiert. Es wurden jeweils die Effekte des Erfassungsbereichs der Erkennung, der Genauigkeit, der Verfügbarkeit und der Zuverlässigkeit der Daten betrachtet.

Ein Vorteil des Ansatzes liegt in der detaillierten Beschreibung der Aspekte (z.B. die TTC) in Abhängigkeit von der Fahrsituation (z.B. von der Entfernung). Damit ist das in der Literatur oft diskutierte Problem der Auswahl eines geeigneten Schwellwertes situationsspezifisch vorangebracht. Ein weiterer Vorteil ist, dass die Interaktionen zwischen dem Ego-Fahrzeug und den Verkehrsteilnehmern berücksichtigt werden, da das SIG auf das eigene Fahrzeug zentriert ist. Viele kartenbasierte Ansätze berücksichtigen lediglich die Position der Objekte. Für diesen Ansatz ist es möglich, dass sich die SIG-Strukturen an einen bestimmten Fahrer, eine Fahrwei-

se oder eine Fahrsituation anpassen. Der Ansatz kann in Abhängigkeit von den verwendeten Trainingsdaten angepasst werden. Auch ein kontinuierliches Anlernen im Verlauf der Nutzung (Detektionsphase) ist hierüber prinzipiell umsetzbar<sup>6</sup>.

Die Nachteile der Ansatzes sind die großen Datenmengen und zum Teil auch die Redundanzen, die im SIG entstehen. Hierzu sollten die am Anfang des Kapitels angesprochenen Optimierungsstrategien für das SIG weiter untersucht werden, um die Daten zu reduzieren und Redundanzen auszuschließen<sup>7</sup>. Ein weiterer Nachteil, den die meisten Arbeiten in der Literatur auch aufweisen, ist, dass die Komponenten der Fahrsituation im Einzelnen betrachtet werden. Für eine Betrachtung der Gesamtsituation wird viel semantische Information benötigt, womit das Datenvolumen exponentiell ansteigen würde. Zum Teil lassen sich jedoch auch semantische Informationen im SIG berücksichtigen, indem die Schwellwerte angepasst werden. Beispielsweise würde in der Nähe einer Schule oder bei starkem Regen der Schwellwert  $\alpha$  höher (bspw. doppelt so hoch) gewählt. Eine letzte Einschränkung des Ansatzes liegt darin, dass nur die momentane Situation betrachtet wird, und nicht Objekttrajektorien oder eine Historie des Geschehens.

---

<sup>6</sup>Dabei können allerdings auch negative Lerneffekte auftreten.

<sup>7</sup>Sind die Unterschiede zwischen den Verteilungen von zwei Zellen im SIG statistisch signifikant, so können sie vereinigt werden (Varianzanalyse).





## 9 Experimentelle Ergebnisse

Die in dieser Arbeit vorgestellten Methoden wurden prototypisch implementiert und getestet. Dazu wurden unterschiedliche Versuchsplattformen verwendet, die in Abschnitt 9.1 beschrieben werden. Nun wird in den weiteren Abschnitten jeweils auf die Experimente zu den drei Hauptebenen der Arbeit eingegangen. Zunächst werden die Methoden für die Bestimmung der Kameraschwankungen und für die Fahrkorridorerkennung diskutiert. Anschließend wird die Objekterkennung erörtert, wobei sowohl die Testszenarien in der simulierten Umgebung als auch die Tests mit realen Daten dargestellt werden. Die Ergebnisse der Objekterkennung werden in drei Konfigurationen aufgeschlüsselt, zunächst jeweils nur mit Kamera und Laserscanner als Einzelsensoren, dann mit Fusion der beiden. Schließlich werden einige Experimente zur Gefahrenerkennung vorgestellt. Dazu wurden simulierte Daten verwendet, womit viele Trainingsdaten und gezielte Gefahrensituationen erzeugt werden konnten.

### 9.1 Plattformen, Sensoren und Testfahrten

Für den experimentellen Einsatz der Methoden dieser Arbeit sind zwei von den in Abschnitt 1.3 eingeführten Plattformen verwendet worden. Zunächst wurde das Versuchsfahrzeug *ViewCar* in einer Testkampagne eingesetzt, in der Autobahn-, Landstraße- und innenstädtischen Szenarien befahren wurden. Der Fokus dieser Experimente liegt vor allem auf Autobahnszenarien. Das *ViewCar* verfügt unter anderem über folgende Sensorsysteme:

- Laserscanner: Der Sensor von der Firma Ibeo GmbH, Modell Alaska XT, scannt mit 12,5 Hz in vier Schichten und führt eine Segmentierung der Objekte durch. Er hat eine ausrichtungsabhängige Winkelauflösung von  $0,125^\circ$  im Frontalbereich und  $0,5^\circ$  in den seitlichen Bereichen.
- Kamera: Die Kamera (CMOS) von der Firma Aglaia GmbH besitzt ein 3,5 mm-Objektiv, eine Pixelgröße von  $7,5 \mu\text{m} \times 7,5 \mu\text{m}$  und eine Bildauflösung von  $640 \times 480$  Pixeln. Es handelt sich um eine Farbkamera, die jedoch für diese Arbeit in Grauwerten betrieben wird. Die Kamera hat eine automatische Helligkeitsanpassung und ist auf dem Armaturenbrett (etwa 50 cm tiefer als der Fahrerblick) montiert, wodurch der Blickwinkel (Nicken) geringer ist, als der des Fahrers.
- Ortungssystem: Das System von der Firma iMAR GmbH integriert ein differentielles GPS mit einer Trägheitsplattform und liefert hochgenaue Positionsdaten. In Verbindung mit einer verfügbaren digitalen Karte werden Informationen über die aktuell befahrene Straße ausgelesen.
- Fahrstreifenerkennung: Es handelt sich um ein videobasiertes System, welches unter anderem Informationen zur Breite, Krümmung, Art und Dicke der Markierungen des befahrenen Fahrstreifens liefert.

Die Sensordaten werden im *ViewCar* aufgezeichnet und mit der genauen Erfassungszeit versehen. Die Verarbeitung der Daten geschieht bisher offline.

In Ergänzung zu den realen Sensoren wurde die Fahrsimulationsumgebung eingesetzt. Der Vorteil dabei ist, dass die realen Daten der Situation bekannt sind (*ground truth*), womit quantitative Aussagen über die Methoden möglich sind. Folgende Sensoren werden realistisch simuliert:

- Der Laserscanner wurde anhand eines Sensormodells simuliert. Darin sind die ausrichtungsabhängige Winkelauflösung sowie die Entfernungsgenauigkeit implementiert. Die gelieferten Objekte beinhalten nur Fahrzeuge und weisen keine Segmentierungsfehler auf.
- Die Kamera wurde anhand der Visualisierungssoftware implementiert, die dem Testfahrer eingeblendet wird. Die damit erhaltenen Bilder besitzen sämtliche Informationen der Fahrsituation. Genauso wie bei realen Bildern ergeben sich hier gelegentlich Artefakte im Bild, die die Bildverarbeitung realistisch erschweren. Die Lichtverhältnisse in den Szenarien sind jedoch konstant.
- Position sowie die Eigenschaften der befahrenen Straße sind direkt aus der Simulation ableitbar.
- Fahrkorridor: Anhand der bekannten Straßendaten und der exakten Position sowie Kurswinkel des Ego-Fahrzeugs werden die Daten des Fahrkorridors erzeugt und für die Objekterkennung zugreifbar, wie sie idealerweise die Fahrkorridorerkennung liefern würde.

Die Simulatorfahrten mit simuliertem Fremdverkehr finden in einem Autobahnszenario statt.

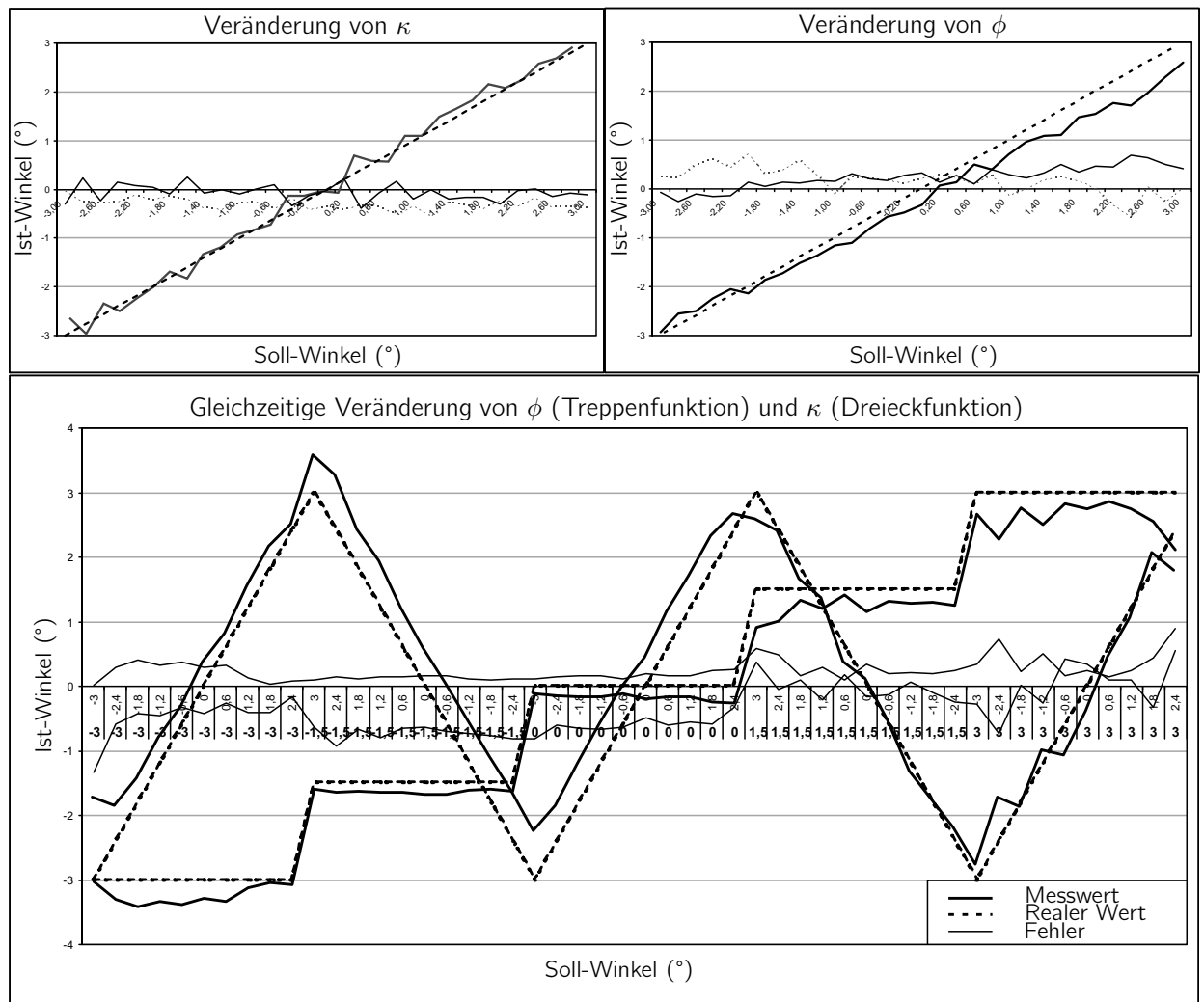
## 9.2 Ergebnisse der Bestimmung der Fahrzeugorientierung

Das in Kapitel 4 vorgestellte Verfahren zur Korrektur von Fahrzeugschwankungen wurde implementiert, auf die Versuchsplattformen angepasst, validiert und getestet.

Zum Validieren der Methode wurde die Ausrichtung der Kamera in einer Situation in der Simulationsumgebung gezielt verändert und mit dem Ergebnis der automatischen Korrektur verglichen. Abbildung 9-1 stellt oben zwei Tests dar, in denen jeweils der Gierwinkel ( $\kappa$  um die  $Z$ -Achse) und der Nickwinkel ( $\phi$  um die  $Y$ -Achse) zwischen  $-3^\circ$  und  $+3^\circ$  gedreht wurden (gestrichelte Kurven). Die durchgezogenen dicken Kurven stellen den ermittelten Wert dar. Der mittlere absolute Fehler<sup>1</sup> (dünne Kurven) der geschätzten Winkel beträgt jeweils  $\overline{\varepsilon_\kappa} = 0,14^\circ$  und  $\overline{\varepsilon_\phi} = 0,27^\circ$ . Im unteren Diagramm wurden beide Winkel gleichzeitig geändert ( $\phi$  als Treppenfunktion und  $\kappa$  als Dreiecksfunktion). Dabei sind die mittleren absoluten Fehler  $\overline{\varepsilon_\kappa} = 0,45^\circ$  und  $\overline{\varepsilon_\phi} = 0,24^\circ$ . Die Korrelation zwischen beiden Fehlern ist mit  $\rho_{\kappa\phi} = 0,41$  relativ hoch, was darauf hinweist, dass die Fehler beider Winkel nicht voneinander unabhängig sind. Trotz dieser Ungenauigkeiten konnte die Methode mit den in Abbildung 9-1 dargestellten Ergebnissen validiert werden, da die realen und die ermittelten Ausrichtungsveränderungen eine große Übereinstimmung zeigen.

Zum weiteren Testen der Methode, wurde diese anhand von realen Testdaten (in Autobahn und Landstraße) bewertet. Um Korrelationen zu vermeiden wurde hier nur der Nickwinkel  $\phi$  geschätzt, da dieser den größten Fehlereinfluss auf die Objekterkennung hat. Da in den Testfahrten keine realen Werte (*ground truth*) zum Vergleich vorlagen, wurde zur Auswertung ein visueller Abgleich durchgeführt. Die Methode weist in einem Autobahnabschnitt bei der verwendeten Kamerakonstellation eine niedrige Verfügbarkeit auf, da sie nur auf 69,75% der aufgezeichneten

<sup>1</sup>Mittlerer absoluter Fehler von  $n$  Schätzungen  $\hat{x}_i$  bei Kenntnis der *ground truth*  $\tilde{x}_i$ :  $\overline{\varepsilon_x} = \frac{1}{n} \sum |\tilde{x}_i - \hat{x}_i|$ .



**Abbildung 9-1:** Validierung der Korrektur von Fahrzeugschwankungen anhand von gesetzten Parametern in der Simulationsumgebung.

Bilder anwendbar ist. In den restlichen Bildern bleiben die Fahrstreifenmarkierungen wegen der niedrigen Bildqualität und der tiefen Einbauposition der Kamera unerkant. Die Quote von groben Fehlern in realen Bildern ist gering (5,11% der aufgezeichneten Bilder). Diese ergeben sich vor allem in Fahrstreifenwechseln oder bei unklaren Fahrstreifenmarkierungen. Da die bestimmten Orientierungsschwankungen von  $\phi$  eine Standardabweichung von  $\sigma_\phi = 0,16^\circ$  besitzen, die in der Größenordnung der in den Simulationsergebnissen festgestellten mittleren Fehler von  $0,24^\circ$  liegt, sind diese Methoden für diese Testfahrt nicht signifikant (die erkannten Schwankungen könnten genau so gut Rauschen sein).

In einem Landstraßenszenario beträgt die Verfügbarkeit 73,88%, die Quote der groben Fehler 6,80% und die Standardabweichung der bestimmten Orientierungsschwankungen  $\sigma_\phi = 0,39^\circ$ . Die höhere Verfügbarkeit hängt mit den dichteren Fahrstreifenmarkierungen zusammen. Sowohl die höhere Standardabweichung als auch die höhere Anzahl von groben Fehlern hängen mit der Tatsache zusammen, dass Landstraßen mehr Unebenheiten als Autobahnen aufweisen. Auch hier ist die Signifikanz der Ergebnisse gering, ein visueller Abgleich zeigt jedoch den richtigen Trend in der Erkennung der Schwankung.

Aus diesen Gründen erweist es sich nicht als sinnvoll, mit der aktuellen Plattformkonfiguration die Methode zur Korrektur der Fahrzeugschwankungen direkt zu verwenden. Da ein zuverlässiger

Trend der Schwankungen erkannt wird (wie in der Validierung zuvor bereits gezeigt), kann mit dieser Methode jedoch die Einstellung weiterer Parameter in der Verarbeitung vorgenommen werden – wie beispielsweise die Erhöhung von Schwellwerten.

Die Methode zeigt vielversprechende Ergebnisse, die eine Alternative zu anderen Methoden aus der Literatur, wie Liang [91] darstellt. Obwohl Liang genauere Ergebnisse erzielt, besitzt die hier vorgestellte Methode zwei wichtige Vorteile, nämlich die Feststellung von ungeeigneten Bildern und die Optimierungssuche der Orientierungsparameter.

### 9.3 Ergebnisse der Fahrkorridorerkennung

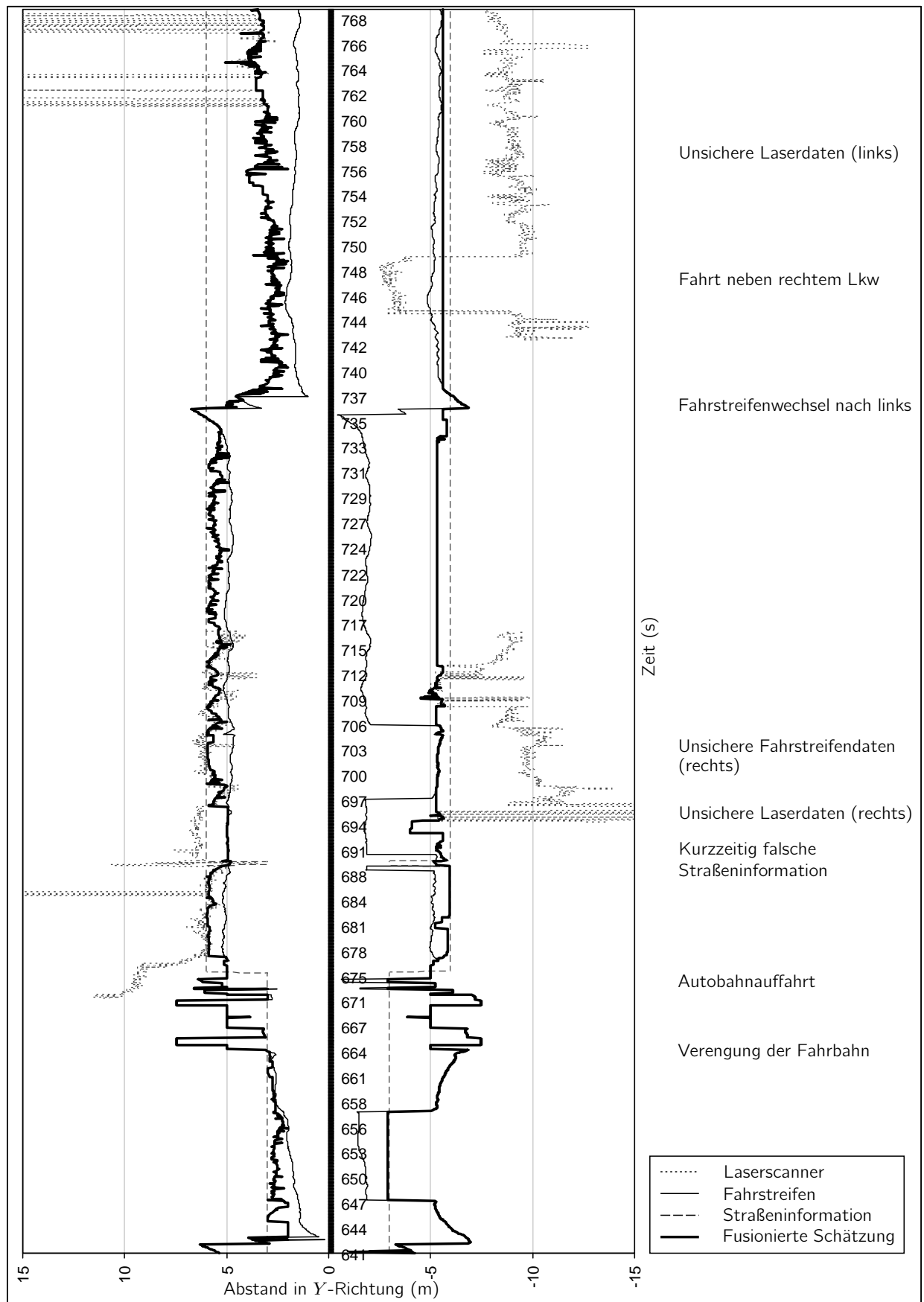
Das in Abschnitt 4.4 entwickelte Verfahren zur Fahrkorridorerkennung wurde prototypisch implementiert und in den durchgeführten Testfahrten für die Unterstützung in der Objekterkennung verwendet. Dazu wurden die Daten aus dem Laserscanner, der im ViewCar befindlichen Fahrstreifenerkennung und die Informationen über die befahrene Straße verwendet.

Abbildung 9-2 stellt beispielhaft die Ergebnisse der Fahrkorridorerkennung in einem 128 sekündigen Abschnitt der Testsequenz dar, die eine Autobahnauffahrt und einen Autobahnabschnitt beinhaltet. Darin sind die Schätzungen aus jeder Einzelquelle sowie das Ergebnis der Fusion zu sehen.

In diesem Sequenzabschnitt läuft die Fahrkorridorerkennung im Ganzen richtig. Nichtsdestotrotz sind darin manche Effekte der Methode ersichtlich, die im Folgenden erläutert werden. In der Sequenz sind mehrere Stellen erkennbar, in denen jeweils die Fahrstreifendaten (wie im Bereich um 647-658 s, 671-675 s, 737 s), die Laserscannerdaten (wie in 656 s, 670 s, 692 s, 743-750 s) oder die digitalen Karten (z.B. 689 s) unsicher sind. Bis auf dem Bereich zwischen 665 s und 673 s ist die Ausgabe der fusionierten Daten richtig. Im besagten Bereich finden gleichzeitig ein Fahrstreifenwechsel und eine Erweiterung der Straße um einen Fahrstreifen statt. Im Zeitpunkt 737 s findet ein Fahrstreifenwechsel statt. Im Zeitbereich zwischen 743 s und 750 s befindet sich das Ego-Fahrzeug neben einem Lkw, wodurch der Laserscanner einen zu schmalen Korridor schätzen würde. Die Fusion liefert jedoch eine richtige Schätzung.

Solche Situationen lassen sich während der gesamten Fahrt erkennen. Die Bestimmung der Fahrkorridor Grenzen ist in den meisten Situationen richtig, was eine effektive Erleichterung für die restliche Datenverarbeitung bedeutet. Mit der momentanen Implementierung ergeben sich allerdings zweierlei Schwierigkeiten in der Fahrkorridorerkennung. Einerseits kommen gelegentlich Situationen vor, in denen der Fahrkorridor aufgrund von engen Kurvenradien oder in Kreuzungsbereichen kurzzeitig zu schmal geschätzt wird. Dies bedeutet eine kurzzeitige Einschränkung der Objekterkennung und damit auch der Gefahrenerkennung. Andererseits wird der Fahrkorridor an manchen Situationen kurzzeitig zu breit geschätzt, wie bspw. bei falschen oder fehlenden Fahrstreifendaten, Daten aus der digitalen Karte oder in Kreuzungsbereichen. Dadurch entstehen auch kurzzeitige Rauschobjekte in der Objekterkennung, die in sehr unwahrscheinlichen Konstellationen zu falschen Gefahreneinschätzungen (*false positive*) führen können.

Aus diesen Experimenten wird ersichtlich, dass eine Verbesserung der Fahrkorridorerkennung insbesondere in den Bereichen mit engen Kurven und Kreuzungen sinnvoll wäre, um eine zuver-



**Abbildung 9-2:** Beispiel der Fahrkorridorerkennung anhand einer Sequenz realer Testdaten.

lässigere Objekt- und Gefahrenerkennung zu ermöglichen.

Im Vergleich mit anderen Arbeiten, wie zum Beispiel Berger et al. [15], Weiß et al. [158], Darms et al. [38] und Loose und Franke [94], lässt sich mit der hier vorgestellten Methode aufgrund der Fusion eine höhere Verfügbarkeit (Ausfallsicherheit) und Zuverlässigkeit erwarten. Im Vergleich zur Arbeit von Loose und Franke [94] ist das hier verwendete Modell einfacher und weniger differenziert (ohne Kurven), benötigt jedoch daher einen geringeren Rechenaufwand.

## 9.4 Ergebnisse der Objekterkennung und -verfolgung

Zum Testen der Objekterkennung und -verfolgung sind zwei Wege verfolgt worden: mit simulierten und mit realen Testdaten.

Als erste Teststrategie sind eine Reihe von gezielten Testszenarien für den Fahrsimulator (mit simulierten Fremdverkehr) vorbereitet und untersucht worden. Damit können quantitative Aussagen über die Methoden getroffen werden. Das Ziel dieser Arbeit ist in erster Linie die Erfassung der Fahrsituation in Autobahnszenarien. Es werden daher folgende drei simulierte Testszenarien dargestellt, die besonders häufig in der Autobahn vorkommen, nämlich eine Fahrgasse (TS1 in Unterabschnitt 9.4.1), Überholen (TS2 in Unterabschnitt 9.4.2) und Fahrstreifenwechsel (TS3 in Unterabschnitt 9.4.3).

Des Weiteren sind Daten aus den in den letzten Abschnitten vorgestellten Testfahrten im realen Fahrzeug betrachtet worden. Daraus wurden zwei repräsentative Realdatenausschnitte (RA) ausgewählt, jeweils in einer dreispurigen Autobahn (RA1 in Unterabschnitt 9.4.4) und beim Überholen (RA2 in Unterabschnitt 9.4.5). Zum Testen der Methoden an realen Fahrtdaten lag keine Information über die *ground truth* vor. Aus diesem Grund wurden für die Ergebnisse ein visueller Vergleich und die geschätzten Genauigkeiten als Kriterien verwendet. Werden vergleichbare Ergebnisse erhalten, wie mit den simulierten Testdaten, so dient das Experiment zur Validierung der Simulation.

Alle Tests sind mittels einer Ergebnistabelle sowie einiger Ablaufdiagramme visualisiert. Die Spalten der Ergebnistabelle sind folgende:

- Die mittlere absolute Abweichung  $\overline{\varepsilon_{x_i}}$  jeder Komponente  $x_i$  (wenn *ground truth* bekannt ist<sup>2</sup>).
- Die mittlere geschätzte Fehlerkovarianz<sup>3</sup>  $\overline{\sigma_{x_i}}$  jeder Komponente  $x_i$ .
- Die Anzahl von Id-Wechseln (*idch*) des verfolgten Fahrzeuges. Eine niedrige Zahl ist auf der Anwendungsebene vorteilhaft.
- Prozentueller Anteil der *false negatives* (*f.n.*, verpasste Objekte) aus den Objekthypothesen und die gesamte Dauer in Sekunden, während der ein Objekt verpasst wurde<sup>4</sup>. Auch der prozentuelle Anteil der Zyklen, in denen ein Objekt in den Rohdaten nicht erfasst

---

<sup>2</sup>Absolute Abweichung bei Kenntnis der *ground truth*:  $\varepsilon_x = \sqrt{(\tilde{x} - \hat{x})^2} = |\tilde{x} - \hat{x}|$ ; Mittlere absolute Abweichung über die Zeit:  $\overline{\varepsilon_x} = \frac{1}{n} \sum \varepsilon_x$ .

<sup>3</sup>Geschätzte Fehlerkovarianz  $\sigma_x^2 = E[(\tilde{x} - \hat{x})(\tilde{x} - \hat{x})^T]$  und mittlerer geschätzter Fehler über die Zeit:  $\overline{\sigma_x} = \frac{1}{n} \sum \sigma_x$ .

<sup>4</sup>Werden zwei Objekte gleichzeitig 1 Sekunde lang nicht verfolgt, so beträgt diese Zahl 2 s.

**Tabelle 9-1:** Quantitative Ergebnisse aus dem Testszenario TS1

Dauer 25 s	$\overline{\varepsilon_x}$ $\overline{\sigma_x}$ (m)	$\overline{\varepsilon_y}$ $\overline{\sigma_y}$ (m)	$\overline{\varepsilon_{\dot{x}}}$ $\overline{\sigma_{\dot{x}}}$ (km/h)	$\overline{\varepsilon_{\dot{y}}}$ $\overline{\sigma_{\dot{y}}}$ (km/h)	$\overline{\varepsilon_w}$ $\overline{\sigma_w}$ (m)	$\overline{\varepsilon_l}$ $\overline{\sigma_l}$ (m)	<i>idch/s</i>	%f.n. (s) %f.n. roh	f.p. id/s s/f.p. f.p.roh(s)
<i>cam</i> 24 Hz	0,891 1,094	0,058 0,257	4,376 36,153	2,669 21,034	0,109 0,402	- -	0,04	0,00 (0,00) 0,21	0,00 0,00 0,42
<i>lsr</i> 12,5 Hz	0,019 0,437	0,019 0,175	0,797 21,546	3,054 14,565	0,064 0,207	- - -	0,00	0,00 (0,00) 0,00	0,00 0,00 0,00
<i>fus</i> 36,5 Hz	0,120 0,512	0,021 0,194	4,193 27,681	2,889 17,201	0,065 0,212	- -	0,00	0,00 (0,00) 16,53	0,28 0,43 5,47

wurde (*f.n. roh*), wird angegeben. Dies zeigt den durch das Verfolgungssystem erreichten Gewinn.

- Die Anzahl von *false positives* (*f.p.*, Phantomobjekten) als Rate von Id pro Sekunde und deren mittlere Dauer. Die dritte Zahl drückt die Anzahl an *false positives* in den Rohdaten (*f.p. roh*) aus und zeigt damit die durch das System erreichte Wegfilterung.

Die Zeilen der Tabelle zeigen jeweils die Ergebnisse anhand der Kamera (*cam*), Laserscanner (*lsr*) und Fusion der beiden (*fus*). Dazu sind jeweils die verfügbaren Datenraten angegeben.

Die Diagramme stellen den Verlauf der *X*-Komponente, der *Y*-Komponente, der Längsgeschwindigkeit ( $\dot{x}$ ) und des NIS-Wertes<sup>5</sup> dar. Letzterer gibt ein Maß zur Stabilität der Verfolgung an (siehe Erläuterungen in Abschnitt 7.4.1).

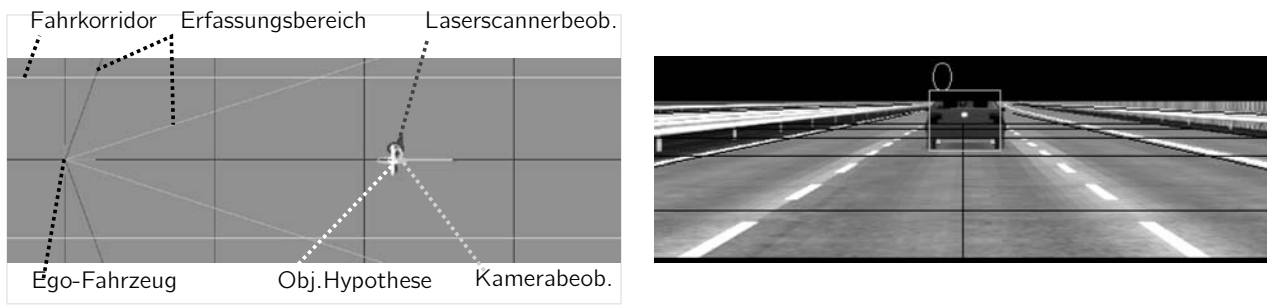
#### 9.4.1 Testszenario TS1: Folgefahrt

Auf einer geraden Autobahn mit drei Fahrstreifen wird auf dem mittleren Fahrstreifen hinter einem führenden Fahrzeug gefahren. Der Abstand zwischen beiden Fahrzeugen ändert sich von 20 m bis etwa 70 m und dann wieder auf 20 m mit einer relativen Geschwindigkeit von ca. 30 km/h. In diesem Testszenario wird die Güte in der Verfolgung in einfachen Situationen getestet. Die Längskomponente rückt dabei in den Vordergrund. Hierzu werden die maximalen Entfernungen, Relativgeschwindigkeiten und Beschleunigungen bzw. Bremsungen getestet, bei denen das System in der Lage ist, die Objekte zu verfolgen.

Die Ergebnisse der Objekterkennung und Verfolgung sind in Tabelle 9-1 und in den Diagrammen in Abbildung 9-4 zusammengefasst. Abbildung 9-3 stellt einen Schnappschuss des Tests dar.

Mithilfe der konkurrierenden und modularen Struktur der Methoden gelingt es dem System, die Fahrsituation anhand jeweils eines Sensors zu verfolgen. Auch in dem Fall, dass die Daten der beiden fusioniert werden, wird die Verfolgung trotz höherer Komplexität erfolgreich durchgeführt.

<sup>5</sup>NIS: *Normalized Innovation Square* oder normierter Innovationsquadrat, siehe Gleichung (7.16) auf Seite 74.



**Abbildung 9-3:** Schnappschuss aus dem Testszenario TS1 zum Zeitpunkt 2 s. Links ist eine orthographische Sicht der Szene und rechts ein Kamerabild mit den Objekthypothesen dargestellt.

Die Genauigkeit der Schätzung (siehe Tabelle 9-1) weist in allen drei Fällen gute Werte auf. Besonders genau erweisen sich die Ergebnisse anhand des Laserscanners, da seine Beobachtungen sehr genau sind und wenige Störeffekte beinhalten. In der fusionierten Version erreicht die Genauigkeit einen zum Teil schlechteren Wert als mit dem Laserscanner. Dies ist der höheren Komplexität und den unsichereren Kamerabeobachtungen zurückzuführen. Dies wird auch im Diagramm zum NIS-Wert sichtbar. Die Genauigkeit der Geschwindigkeit ist sehr niedrig. Dies erklärt sich durch den Initialisierungsfehler (zwischen 0 s und 2 s). Dank einer Objektverdoppelung kann jedoch dabei das Objekt mit derselben Id verfolgt werden, was auf der Anwendungsebene vorteilhaft ist. Die Länge des Objektes ist in diesem Experiment nicht beobachtbar (siehe Abbildung 9-3).

In allen drei Konfigurationen werden hohe Zuverlässigkeitswerte erreicht. Im Fall der Fusion erscheinen wenige Phantomobjekte (*f.p.*), die jeweils sehr kurze Lebensdauer (0,28 s) und damit keine wichtigen Auswirkungen aufweisen. Die Zahl der verpassten Objekte (*f.n.*) in den Rohdaten bei der Fusion bezieht sich auf die Zeit, in der sich das Objekt außer Reichweite der Kamera befindet.

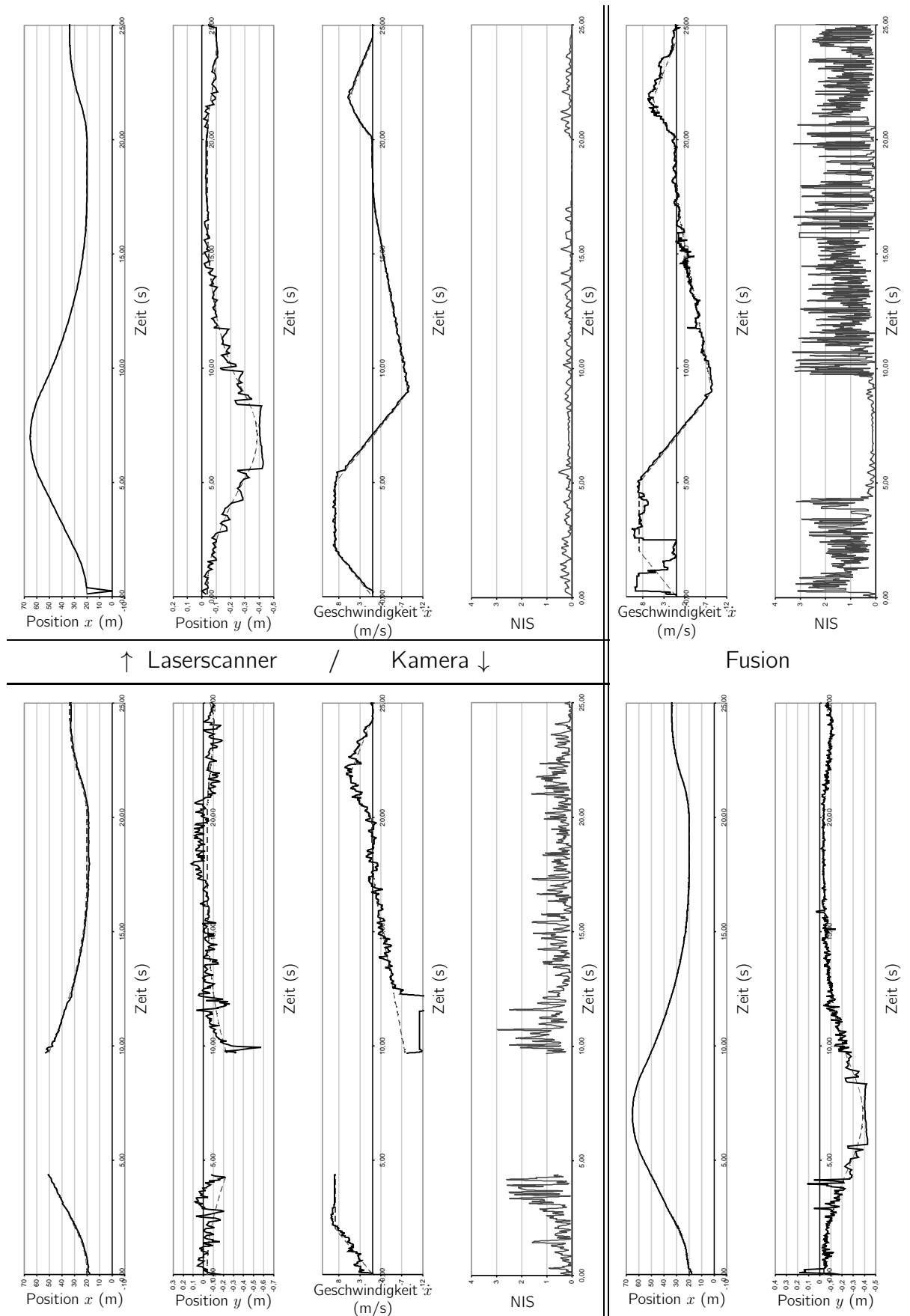
Die Fusion ermöglicht eine deutliche Erhöhung der Datenverfügbarkeit, für die eine leicht geringere Genauigkeit in Kauf genommen wird. Die Reichweite der Sensoren wird erfolgreich zusammengefügt, auch wenn nicht im gesamten Bereich fusioniert wird (z.B. in vom Sensor weit entferntem Bereich).

Aufgrund der höheren Verfügbarkeit sind höhere Relativgeschwindigkeiten erkennbar, auf die das System hin untersucht wurde. Bei der aktuellen Konfiguration ist es dem Laserscanner nicht möglich Relativgeschwindigkeiten über 70 km/h zu verfolgen, die beispielsweise bei Gegenverkehr stattfinden können. Die Bildverarbeitung (angenommen, dass genügend Rechenleistung zur Verfügung steht) kann bis ca. 200 km/h verfolgen, jedoch mit eingeschränkter Genauigkeit. Die Fusion zeigt in solchen Extremfällen (z.B. bei Gegenverkehr) ihre Vorteile, denn die Verfolgung kann bei höherer Geschwindigkeit mit besserer Genauigkeit durchgeführt werden.

### 9.4.2 Testszenario TS2: Überholen

Auf einer geraden Autobahn mit drei Fahrstreifen wird auf dem mittleren Fahrstreifen hinter einem führenden Fahrzeug mit variierendem Abstand gefahren, während die Lkws auf dem rechten Fahrstreifen überholt werden und die Pkws auf dem linken Fahrstreifen überholen. Alle Geschwindigkeiten zeigen leichte Veränderungen. Mit diesem Testszenario wird die Fähigkeit des Systems

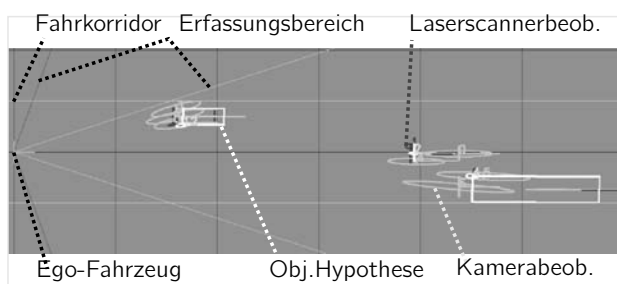




**Abbildung 9-4:** Ablauf des Testszenarios TS1 anhand der Kamera, des Laserscanners und der Fusion. Die Diagramme beinhalten jeweils die Position in X- und Y-Richtung, die Längsgeschwindigkeit und den NIS-Wert. Jedes Kurvensegment entspricht einer Objektdetektion. Die gestrichelten Kurven stellen die *ground truth* dar.

**Tabelle 9-2:** Quantitative Ergebnisse aus dem Testszenario TS2

Dauer 20 s	$\overline{\varepsilon_x}$ $\overline{\sigma_x}$ (m)	$\overline{\varepsilon_y}$ $\overline{\sigma_y}$ (m)	$\overline{\varepsilon_{\dot{x}}}$ $\overline{\sigma_{\dot{x}}}$ (km/h)	$\overline{\varepsilon_{\dot{y}}}$ $\overline{\sigma_{\dot{y}}}$ (km/h)	$\overline{\varepsilon_w}$ $\overline{\sigma_w}$ (m)	$\overline{\varepsilon_l}$ $\overline{\sigma_l}$ (m)	<i>idch/s</i>	%f.n. (s) %f.n. roh	f.p. id/s s/f.p. f.p.roh(s)
<i>cam</i> 24 Hz	2,608 1,078	0,689 0,221	10,639 28,863	5,683 12,619	0,358 0,361	3,594 1,051	0,05	12,32 (6,19) 15,38	0,70 0,44 15,97
<i>lsr</i> 12,5 Hz	0,142 0,566	0,048 0,308	3,287 31,038	5,575 23,959	0,106 0,279	0,207 0,982	0,10	0,00 (0,00) 0,03	0,20 0,68 6,93
<i>fus</i> 36,5 Hz	0,319 0,583	0,068 0,271	3,31 7,56	4,71 3,55	0,120 0,257	0,408 0,507	0,22	0,00 (0,00) 23,68	1,77 0,40 19,33

**Abbildung 9-5:** Schnappschuss aus dem Testszenario TS2 zum Zeitpunkt 3 s. Links ist eine orthographische Sicht der Szene und rechts ein Kamerabild mit den Objekthypothesen dargestellt.

geprüft, mit komplexeren Situationen umzugehen. Die Herausforderung besteht vor allem darin, die Objekte einzeln zu verfolgen, auch wenn sie zum Teil mit sehr kurzem Abstand fahren – sowohl in Längs- als auch in Querrichtung.

Dieses Szenario wurde wie das Vorige in drei Konfigurationen durchgeführt. Die Ergebnisse sind analog in Tabelle 9-2 und den Abbildungen 9-5 und 9-6 dargestellt.

In diesem Testszenario haben alle drei Experimente erfolgreiche Ergebnisse gezeigt. Die Anmerkungen zur Genauigkeit und Zuverlässigkeit aus dem letzten Testszenario sind auch bei diesem vergleichbar. Zusätzlich werden folgende Eigenschaften beobachtet:

Insgesamt ist die Genauigkeit des Laserscanners höher und die Reichweite größer als bei der Kamera. Durch Fusion der beiden wird zusätzlich eine höhere Datenverfügbarkeit erreicht. Der Laserscanner hat beispielsweise in der Sequenz zwischen 2 s und 4 s Schwierigkeiten ein Objekt zu initialisieren. Zum Teil zieht in den Ergebnissen die höhere Komplexität durch die Fusion beider Sensoren eine Verschlechterung der Genauigkeit nach sich – analog zum letzten Experiment.

Die Kamera hat im Vergleich zum Laserscanner einen eingeschränkten horizontalen Erfassungsbereich. Da die Methoden der Bildverarbeitung bezogen auf die verwendeten Kameras ihre Grenzen ab einer größeren Entfernung als 50 m aufweisen, ist die Güte der Daten auf diesen Bereich beschränkt. Die Bildverarbeitung hat ihre Funktion bei unterschiedlichen Objekttypen gezeigt. Dabei wird jedoch deutlich, dass sich beide Sensoren zum Teil auf unterschiedliche Objektteile für die Positionsbestimmung beziehen. Zum Beispiel kann bei einem Bus die hintere Unterkannte durch undeutlichen Schatten ungenau ermittelt werden. Aufgrund dieses systematischen Fehlers, der in der Fusion schwer auszufiltern ist, werden zum Teil zwei Objekte statt einem verfolgt (siehe Ungenauigkeiten in der Fusion der  $X$ -Position in Diagramm 9-6 zwischen 14 s und 16 s).

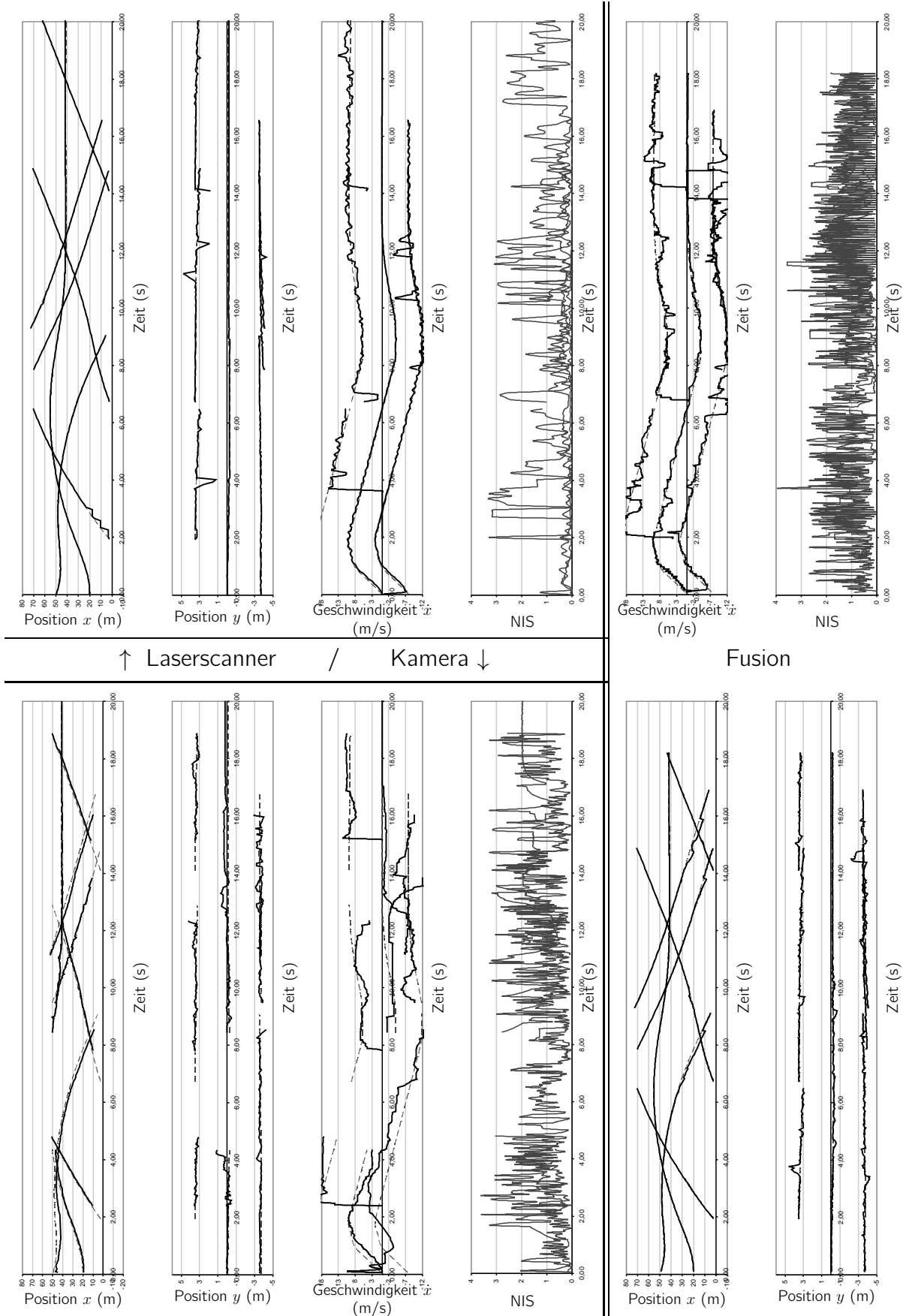
Die Anmerkungen zur Zuverlässigkeit (*f.n.* und *f.p.*) sind wie im letzten Testszenario. Dabei kommt es zum Teil zu Teilverdeckungen zwischen den Objekten, deren Effekte durch die Fusion verringert werden.

In diesem Szenario wird auch die Unterscheidbarkeit von sehr nah beieinander fahrenden Fahrzeugen durch das System deutlich. Diese Tests waren mit seitlichen Abständen von ca. 2 m und Längsabständen von unter 10 m erfolgreich.

### 9.4.3 Testszenario TS3: Fahrstreifenwechsel

Auf einer geraden Autobahn mit drei Fahrstreifen wird auf dem mittleren Fahrstreifen gefahren. Auf dem rechten Fahrstreifen fahren langsamere Fahrzeuge. Ein 10 km/h schneller fahrendes Fahrzeug überholt das Ego-Fahrzeug von links (Zeitpunkt 5 s), wechselt auf den mittleren Fahrstreifen etwa 40 m vor dem Ego-Fahrzeug (Zeitpunkt 7 s). Dabei wechselt das Ego-Fahrzeug auf den linken Fahrstreifen. Schließlich wechselt das Ego-Fahrzeug auf den mittleren Fahrstreifen (Zeitpunkt 12 s), während ein Fahrzeug aus dem rechten Fahrstreifen beschleunigt und zwei Fahrstreifen nach links wechselt (Zeitpunkte 15 s und 20 s).

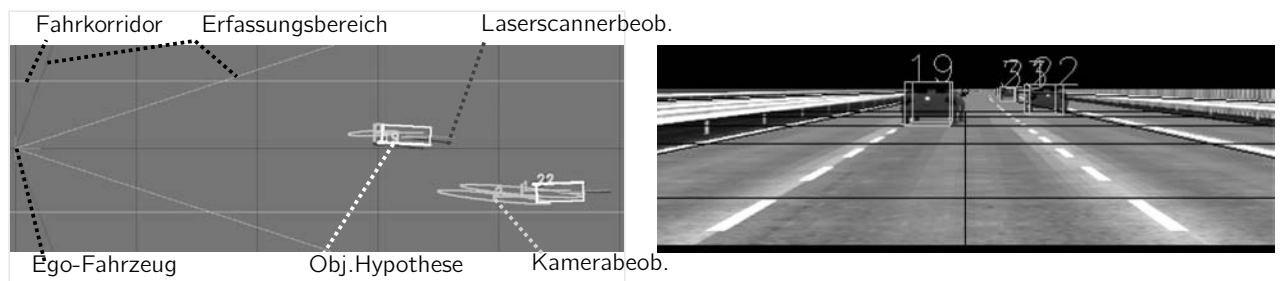
In diesem Testszenario wird insbesondere die Fähigkeit des Systems getestet, die seitlichen Bewegungen der Objekte zu schätzen und zu verfolgen. Zusätzlich werden die Querbewegungen



**Abbildung 9-6:** Ablauf des Testszenarios TS2 anhand der Kamera, des Laserscanners und der Fusion. Die Diagramme beinhalten jeweils die Position in X- und Y-Richtung, die Längsgeschwindigkeit und den NIS-Wert. Jedes Kurvensegment entspricht einer Objektdetektion. Die gestrichelten Kurven stellen die *ground truth* dar.

**Tabelle 9-3:** Quantitative Ergebnisse aus dem Testszenario TS3

Dauer 19 s	$\overline{\varepsilon_x}$ $\overline{\sigma_x}$ (m)	$\overline{\varepsilon_y}$ $\overline{\sigma_y}$ (m)	$\overline{\varepsilon_{\dot{x}}}$ $\overline{\sigma_{\dot{x}}}$ (km/h)	$\overline{\varepsilon_{\dot{y}}}$ $\overline{\sigma_{\dot{y}}}$ (km/h)	$\overline{\varepsilon_w}$ $\overline{\sigma_w}$ (m)	$\overline{\varepsilon_l}$ $\overline{\sigma_l}$ (m)	<i>idch/s</i>	%f.n. (s) %f.n. roh	f.p. id/s s/f.p. f.p.roh(s)
<i>cam</i> 24 Hz	0,492 1,706	0,205 0,306	4,53 5,33	4,24 3,38	0,156 0,493	1,126 1,439	0,14	4,61 (1,16) 11,18	0,50 0,37 8,01
<i>lsr</i> 12,5 Hz	0,059 0,556	0,039 0,312	5,481 31,236	6,702 20,848	0,060 0,267	0,112 0,723	0,09	5,79 (2,38) 10,04	0,27 0,15 3,73
<i>fus</i> 36,5 Hz	0,173 0,665	0,109 0,311	5,620 31,581	7,548 19,529	0,091 0,289	0,242 0,535	0,18	3,72 (1,30) 29,26	1,14 0,19 13,12

**Abbildung 9-7:** Schnappschuss aus dem Testszenario TS3 zum Zeitpunkt 6 s. Links ist eine orthographische Sicht der Szene und rechts ein Kamerabild mit den Objekthypothesen dargestellt.

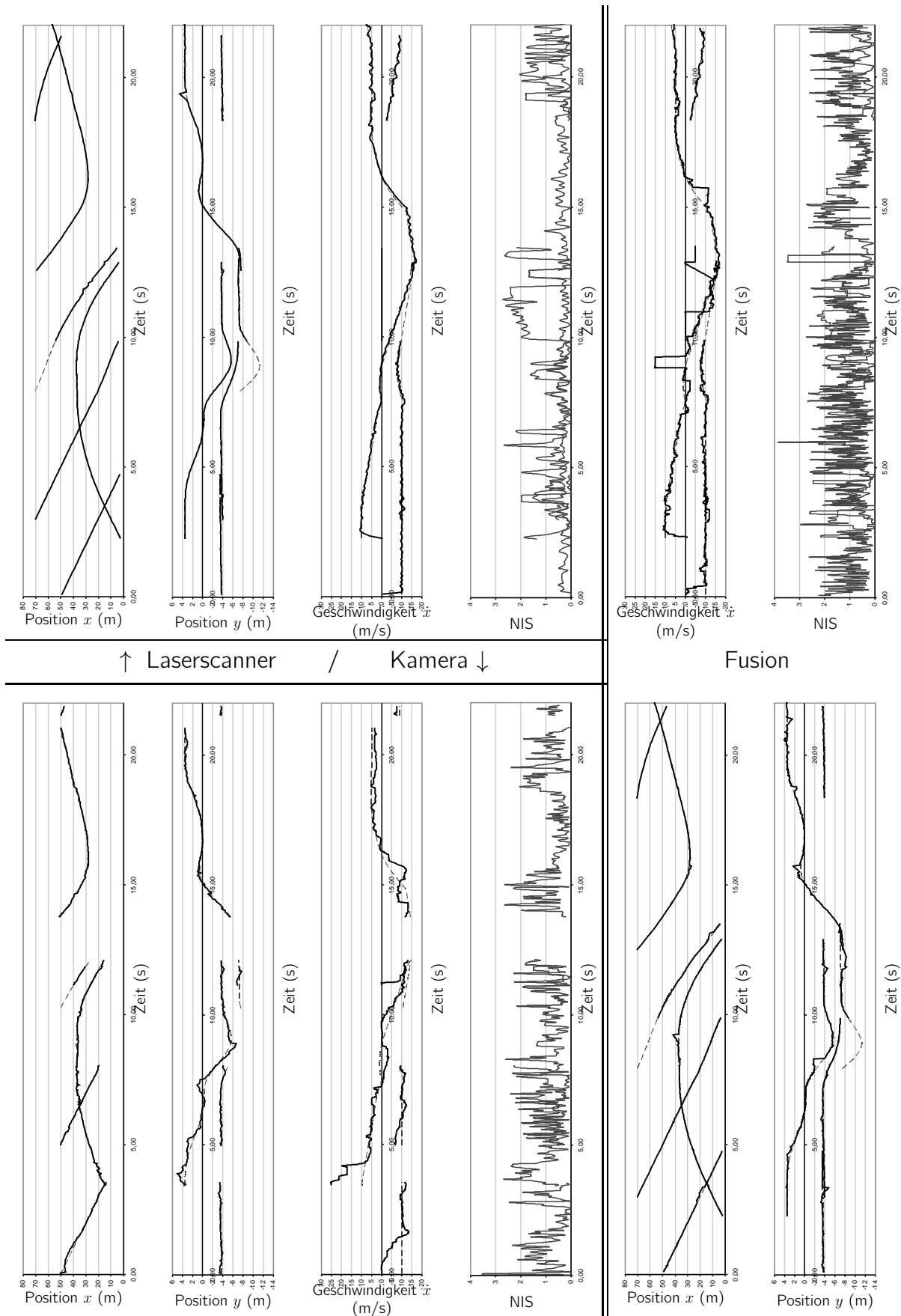
des eigenen Fahrzeugs getestet. Es wird hiermit auch die Verwendung des Fahrkorridors getestet.

Die Ergebnisse dieses Tests sind analog zu den vorigen Resultaten in Tabelle 9-3 und in den Abbildungen 9-7 und 9-8 dargestellt.

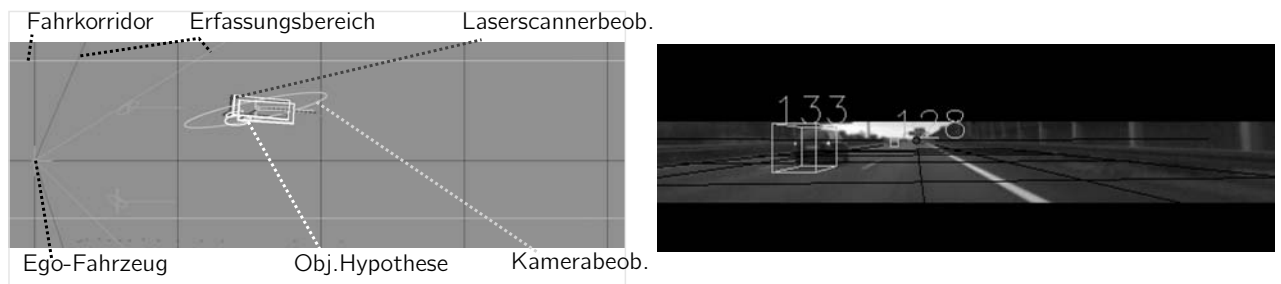
Darin ist zu erkennen, dass es zu einer Objektverdeckung um den Zeitpunkt 10 s kommt. Daher ist die Quote von verpassten Objekten (*f.n.*) höher als bei den vorigen Experimenten.

Es kommt in dem Testszenario zu mehreren Fahrstreifenwechseln. In dem Testszenario werden alle Fahrstreifenwechsel richtig verfolgt (ohne Wechsel der Objekt-Id). Die in Tabelle 9-3 angegebenen Id-Wechsel passieren im Zusammenhang mit der Initialisierung oder Löschung von Objekthypothesen. Bei weiteren Tests, in der die Quergeschwindigkeit der Objekte erhöht wurde, hat sich wieder zeigt, dass der Laserscanner eine höhere Zuverlässigkeit besitzt als die Kamera, diese jedoch eine höhere Datenverfügbarkeit aufweist. Deswegen werden die Fahrstreifenwechsel durch den Laserscanner (wegen zu großer Sprünge) und durch die Kamera (wegen Ausreißen) zum Teil nicht richtig verfolgt. Durch die Fusion und insbesondere aufgrund der Verdoppelungsstrategie erhalten diese Experimente bessere Ergebnisse. Eine fehlerfreie Objektverfolgung beim Fahrstreifenwechseln mit erhöhter Quergeschwindigkeit wird jedoch nicht immer erreicht.

In weiteren Tests ist auch die Empfindlichkeit des Systems auf eigene Querbewegungen festgestellt worden. Durch Veränderung des eigenen Kurswinkels weisen alle Objekte in der Fahrsi-



**Abbildung 9-8:** Ablauf des Testszenarios TS3 anhand der Kamera, des Laserscanners und der Fusion. Die Diagramme beinhalten jeweils die Position in X- und Y-Richtung, die Längsgeschwindigkeit und den NIS-Wert. Jedes Kurvensegment entspricht einer Objektdetektion. Die gestrichelten Kurven stellen die *ground truth* dar.



**Abbildung 9-9:** Schnappschuss aus dem Realdatenausschnitt RA1 zum Zeitpunkt 523,5 s. Links ist eine orthographische Sicht der Szene und rechts ein Kamerabild mit den Objekthypothesen dargestellt.

tuation plötzlich erhöhte Quergeschwindigkeiten auf, die proportional zu deren Entfernung sind. Aus diesem Grund scheint eine wichtige Verbesserung des Systems zu sein, die eigene Bewegung in die Filterung einzubeziehen, um die Objekte in globalen Koordinaten zu verfolgen.

#### 9.4.4 Realdatenausschnitt RA1: Überholen in dreispuriger Autobahn

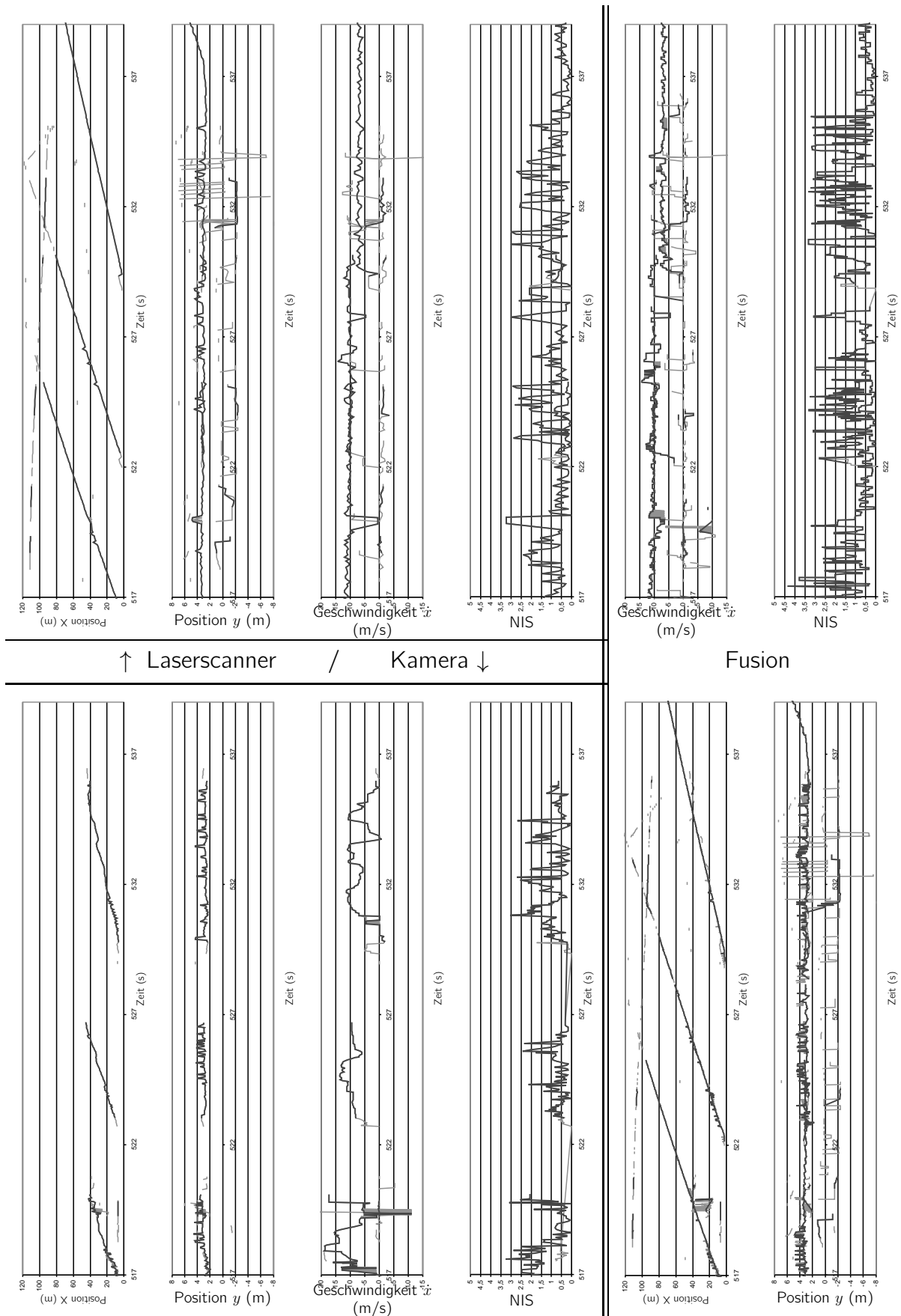
Der erste Ausschnitt aus realen Daten ist in einer dreispurigen Autobahn (auf der A2 bei Braunschweig-Flughafen) aufgezeichnet worden. Dabei wird das Ego-Fahrzeug von mehreren Fahrzeugen von links überholt. In 100 m Entfernung fährt ein Lkw, der nur selten durch den Laserscanner erfasst wird. Das Interesse dieses Tests liegt in der Initialisierung immer neuer Objekte und Verfolgung bei steigender Entfernung.

Tabelle 9-4 zeigt die Ergebnisse der Verarbeitung. Abbildung 9-9 zeigt einen Schnappschuss des Tests. In Abbildung 9-10 sind die Ablaufdiagramme des Tests für jeden Sensor im Einzelnen und für die Fusion der beiden dargestellt.

Dieses Experiment zeigt die Tauglichkeit der Methoden für reale Messdaten. Die Kamerabilder sind sehr kontrastarm und weisen einen sehr geringen Blickwinkel auf die Verkehrsszene auf. Deswegen ist deren Erfassungsbereich praktisch auf 45 m begrenzt. Auch wenn der Laserscanner einen deutlich größeren Erfassungsbereich besitzt, sind Objekte in größeren Entfernungen (ab 80 m) auch schwer zu verfolgen. Innerhalb der erwähnten Erfassungsbereiche läuft die Objekterkennung und Verfolgung erfolgreich. Die geschätzten Genauigkeiten (Fehlerkovarianzen) liegen in einem ähnlichen Bereich wie in den Simulatortests, welches als Validierung der simulierten Sensordaten gilt. Nach visuellem Abgleich werden keine Objekte der Szene verpasst, auch wenn

**Tabelle 9-4:** Quantitative Ergebnisse aus dem Realdatenausschnitt RA1.

Dauer 22 s	$\overline{\sigma_x}$ (m)	$\overline{\sigma_y}$ (m)	$\overline{\sigma_{\dot{x}}}$ (km/h)	$\overline{\sigma_{\dot{y}}}$ (km/h)	$\overline{\sigma_w}$ (m)	$\overline{\sigma_l}$ (m)	idch/s
<i>cam</i> 24 Hz	1,689	0,442	47,269	24,211	0,792	1,607	0,00
<i>lsr</i> 12,5 Hz	0,442	0,266	24,145	19,034	0,475	2,818	0,00
<i>fus</i> 36,5 Hz	0,565	0,386	26,118	21,454	0,539	3,075	0,40



**Abbildung 9-10:** Ablauf des Realdatenausschnitts RA1 anhand der Kamera, des Laserscanners und der Fusion. Die Diagramme beinhalten jeweils die Position in X- und Y-Richtung, die Längsgeschwindigkeit und den NIS-Wert. Jedes Kurvensegment entspricht einer Objektdetektion. Die grauen Kurven entsprechen den Objekten mit niedriger Zuverlässigkeit.



**Tabelle 9-5:** Quantitative Ergebnisse aus dem Realdatenausschnitt RA2.

Dauer 24 s	$\overline{\sigma_x}$ (m)	$\overline{\sigma_y}$ (m)	$\overline{\sigma_{\dot{x}}}$ (km/h)	$\overline{\sigma_{\dot{y}}}$ (km/h)	$\overline{\sigma_w}$ (m)	$\overline{\sigma_l}$ (m)	<i>idch/s</i>
<i>cam</i> 24 Hz	0,579	0,202	15,393	9,820	0,448	2,939	0,00
<i>lsr</i> 12,5 Hz	0,157	0,171	16,990	18,161	0,322	1,149	0,00
<i>fus</i> 36,5 Hz	0,357	0,202	12,499	10,329	0,155	0,296	0,00

kurzzeitige Ausfälle in den Rohdaten vorkommen. Die beiden Sensoren besitzen keine störenden *false positives*. Im seitlichen Bereich erscheinen gelegentlich welche, die jedoch mit realen Objekten (Leitplanken, Büschen) zusammenhängen.

#### 9.4.5 Realdatenausschnitt RA2: Folgefahrt in zweispuriger Autobahn

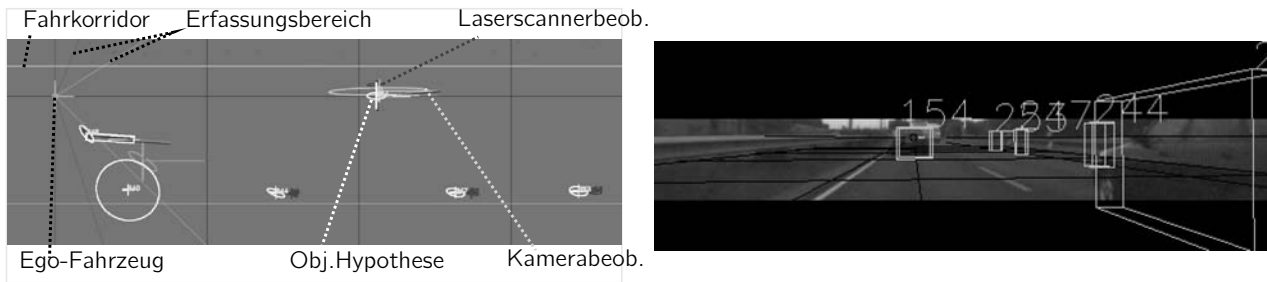
Der zweite Ausschnitt aus realen Daten ist aus einer Fahrt auf einer zweispurigen Autobahn (A391 in Braunschweig) abgeleitet. Auf der rechten Seite ist ein Bereich mit Leitbaken abgetrennt. Es wird hinter einem Vordermann mit variierender Entfernung gefahren. Das Ego-Fahrzeug wird dabei von einem anderen Fahrzeug langsam rechts überholt, das später von der Autobahn abfährt.

Die Ergebnisse des Experiments sind der Tabelle 9-5 und den Abbildungen 9-11 und 9-12 zu entnehmen.

In allen drei Konfigurationen ist die Objekterkennung und Verfolgung trotz des geringeren Erfassungsbereichs der Kamera gelungen. Der Laserscanner ist in der Lage schmalere Objekte (die Leitbaken) zu erkennen und zu verfolgen (siehe Streifen im *X*-Diagramm und Messwerte bei -6 m in *Y*-Diagramm). Da diese Objekte statisch sind, wird die eigene Geschwindigkeit aus dem dritten Diagramm deutlich (etwa 90 km/h). Im Zeitpunkt 790 s fährt ein weiteres Fahrzeug von rechts ins Blickfeld des Laserscanners. Dieses wird erst später (bei 798 s) von der Kamera erkannt. In einer Entfernung von 50 bis 60 m wird immer wieder ein Lkw erkannt. Wegen der geringen Datenverfügbarkeit kann er nicht dauerhaft verfolgt werden. Auch in diesem Datenabschnitt treten Phantomobjekte in den Seitenbereichen auf, die für die Hauptobjekte keine weiteren negativen Effekte haben.

#### 9.4.6 Zusammenfassung der Ergebnisse zur Objekterkennung

Unter Verwendung der in den Kapiteln 5 bis 7 vorgestellten Strategien ist die Erkennung und Verfolgung der Objekte in den Experimenten gelungen. Dabei zeigt sich eine hohe Robustheit gegenüber ungenauen Daten und Ausreißern. Die Methoden funktionieren sowohl für einen einzelnen als auch für mehrere Sensoren. Sie wurden sowohl anhand von gezielten realistisch simulierten Autobahnsituationen, als auch mit vergleichbaren Realdatenabschnitten getestet. Zu den simulierten Szenarien gehören Folgefahrten, Überholmanöver und Fahrstreifenwechseln. Darin sind zum Teil durch das Modell nicht verfolgbare Manöver enthalten (zum Beispiel in



**Abbildung 9-11:** Schnappschuss aus dem Realdatenausschnitt RA2 zum Zeitpunkt 795s. Links ist eine orthographische Sicht der Szene und rechts ein Kamerabild mit den Objekthypothesen dargestellt.

den Testszenarien TS1 mit Beschleunigung und Bremsung und TS3 Fahrstreifenwechsel), die jedoch durch das System korrekt verfolgt werden konnten. Auch Situationen kommen vor, in denen zwei Objekte sehr nah beieinander fahren, so dass das System sie trennen muss (wie in Testszenario TS2). Die Testergebnisse in den simulierten Szenarien zeigen hohe Übereinstimmung mit der bekannten *ground truth*. Bei Realdaten zeigt ein visueller Abgleich ebenso gute Ergebnisse bei ähnlichen Schätzgenauigkeiten. Damit sind die erkannten Objektdaten für die nächste Abstraktionsebene, die Gefahrenerkennung, geeignet.

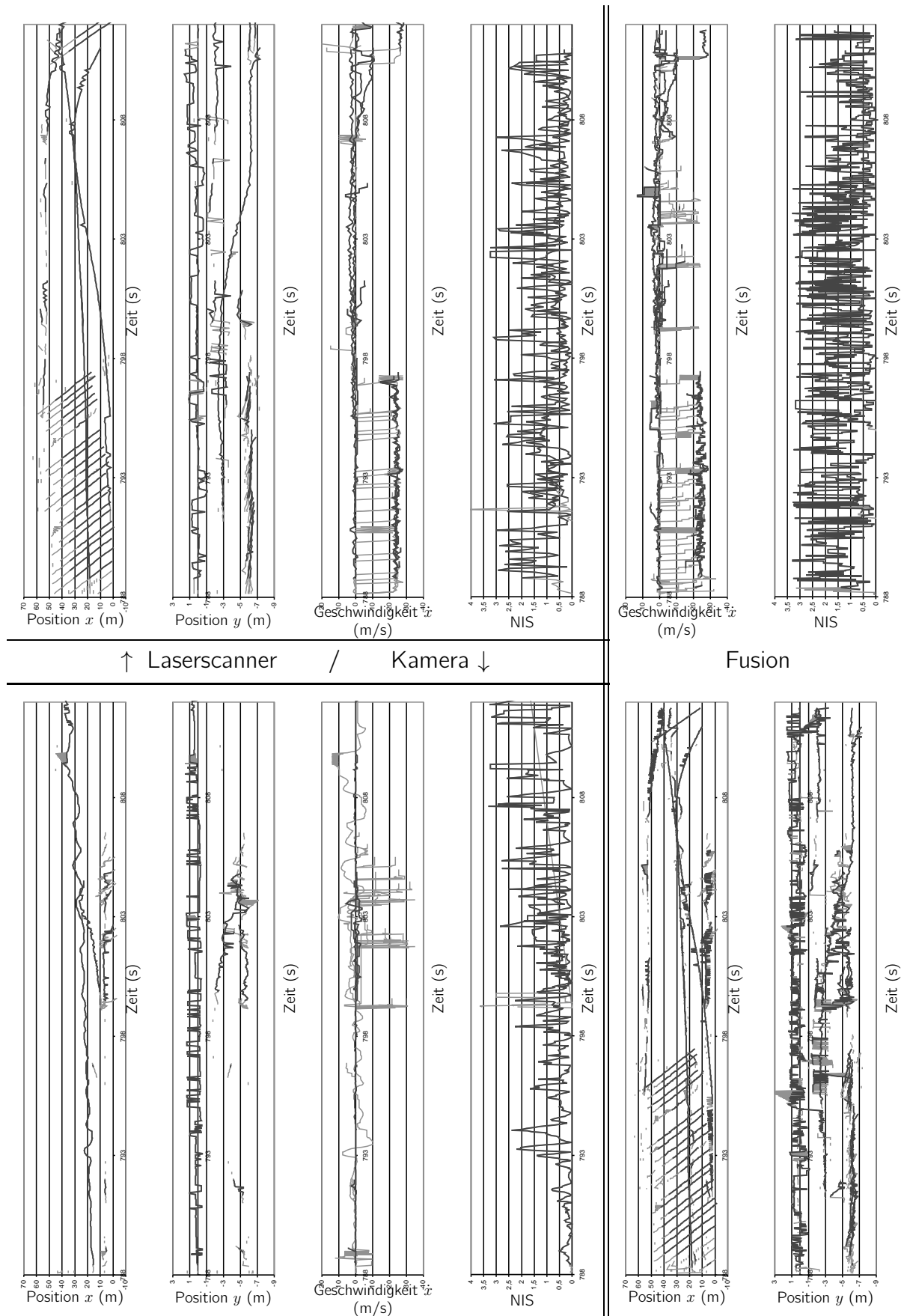
Anhand der ausführlichen Analyse sind einige Schwachpunkte des Systems festgestellt worden. Ein Beispiel ist der negative Einfluss, den die unsichere Erkennung aus den Kamerabildern auf die Fusion erzeugen kann (leicht schlechtere Genauigkeit und Zuverlässigkeit durch höhere Komplexität). Problematische Situationen entstehen zum Beispiel bei Durchfahrten unter eine Brücke oder bei lateralen Bewegungen der Objekte gezeigt. Insbesondere die eigene Kurswinkeländerung hat negative Folgen auf die Verfolgung der Daten.

Im Zusammenhang mit der Fusion sind in der Objekterkennung und Verfolgung deutliche Vorteile zu verzeichnen. Aufgrund der Verwendung mehrerer Sensoren in einer konkurrierenden Art und Weise erhält das Fusionssystem eine hohe Sensorunabhängigkeit, Erweiterbarkeit und niedrige Ausfallsrate. Des Weiteren ist die gemeinsame Verfügbarkeit des Fusionssystems auch höher als bei jedem Einzelsensor, womit Objekte mit höheren Relativgeschwindigkeiten oder Manövern auf komplementärer Art verfolgt werden können. Auch komplementär ist der Gewinn an Beobachtbarkeit durch die Fusion, wodurch unterschiedliche Objektperspektiven erfolgreich zusammengeführt wurden.

Bei der Durchführung der Experimente wurde Wert auf die zuverlässige Funktion und die Qualität der Ergebnisse gelegt. Die Methoden laufen daher zur intensiven Analyse zurzeit noch nicht in Echtzeit.

Ein Vergleich mit anderen Arbeiten aus der Literatur ist aufgrund der unterschiedlichen Sensor-konfigurationen und Testdaten nur schwer möglich. Nichtsdestotrotz können folgende wichtige Punkte hervorgehoben werden.

Im Vergleich zu Systemen, die auf einem einzelnen Sensor basieren oder eine Fusion auf sensornaher Ebene durchführen, ist die Zuverlässigkeit und Erweiterbarkeit dieses Ansatzes eine Stärke, der Rechenaufwand und zum Teil die Genauigkeit jedoch eine Schwäche.



**Abbildung 9-12:** Ablauf des Realdatenausschnitts RA2 anhand der Kamera, des Laserscanners und der Fusion. Die Diagramme beinhalten jeweils die Position in X- und Y-Richtung, die Längsgeschwindigkeit und den NIS-Wert. Jedes Kurvensegment entspricht einer Objektdetektion. Die grauen Kurven entsprechen den Objekten mit niedriger Zuverlässigkeit.

In der Bildverarbeitung basieren andere Arbeiten, wie von Holt [73] und Cramer [35], auf bestimmten Merkmalen, von denen eine korrekte Objekterkennung abhängt. In dieser Arbeit stellt die Objektextraktion basierend auf Texturmerkmalen dagegen eine allgemeine Methode dar.

Im Bezug auf das Dynamikmodell verwenden andere Autoren, wie Cramer [35], komplexe Modelle, können jedoch unbekannte Manöver nicht verfolgen. Becker [13] verwendet dazu eine angepasste Messunsicherheit im Falle eines Manövers. Kämpchen, Fürstenberg und Dietmayer [80, 79] verwenden ein IMM-Filter. Das Dynamikmodell des in dieser Arbeit vorgestellten Ansatzes ist dagegen sehr einfach und schlank, womit die Filterung leicht initialisiert und stabil verfolgt werden kann. Durch die zusätzlichen Strategien werden Objekte auch bei unbekannten Manövern gut verfolgt.

Nur wenige Arbeiten in der Literatur beschäftigen sich mit der wechselnden Perspektive und der partiellen Beobachtbarkeit der Objekte. Diese Aspekte berücksichtigt Stüker [144], betrachtet dabei jedoch nicht explizit die wechselnde logische Referenz eines Objekts, womit die Assoziation auch erschwert wird. Die Konzepte der Beobachtbarkeit wurden in dieser Arbeit weiterentwickelt und explizit für jede Zustands- und Beobachtungskomponente verwendet.

Auch eine selten angesprochene Problematik stellen das *split* und das *merge* von Objekten dar. Hierzu sind in dieser Arbeit Strategien vorgestellt und getestet worden, die damit umgehen können.

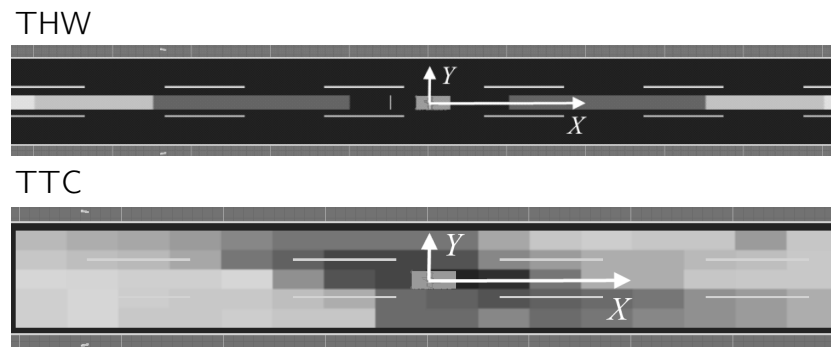
Wie anfangs bereits erwähnt, besitzen manche Arbeiten in der Literatur eine höhere Objektgenauigkeit und Zuverlässigkeit, als die hier vorgestellte. Nichtsdestotrotz findet sich im hier vorgestellten Ansatz eine Sammlung von Strategien, die eine größere Vielfalt von Schwierigkeiten lösen. Damit zeichnet sich der Ansatz in vielen Situationen durch seine hohe Robustheit aus.

## 9.5 Ergebnisse der Gefahrenerkennung

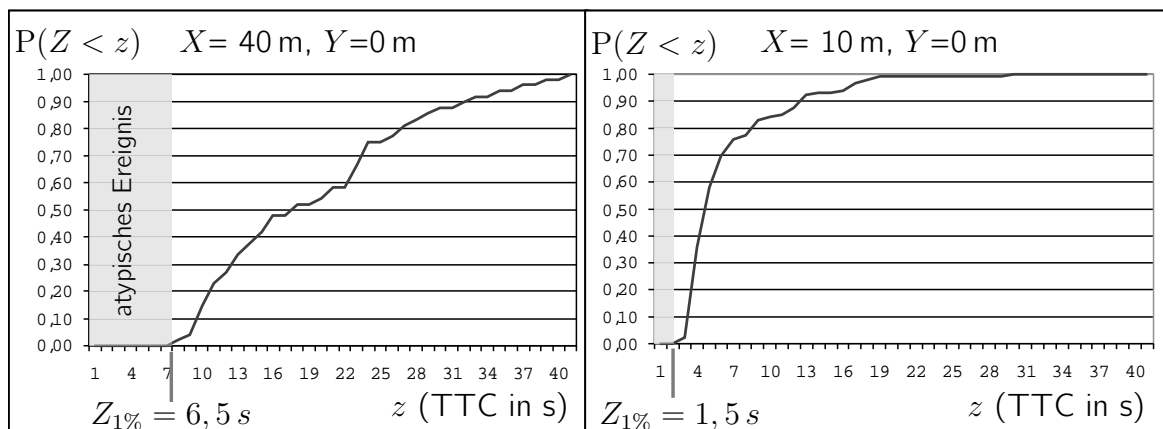
Das in Kapitel 8 vorgestellte Konzept des SIG wurde für die TTC und die THW prototypisch implementiert und mit simulierten Daten getestet.

Die SIG-Struktur wurde in Form zweier Schichten mit positionsabhängigen Häufigkeitsverteilungen definiert, eine für TTC und eine für THW. Da die simulierten Daten lediglich auf geraden Strecken basieren, entsprechen die kartesischen Koordinaten den straßengeführten Koordinaten. Die Längsrichtung des Gitters wurde mit einer Auflösung von 5 m und die Querrichtung mit einer Auflösung von 3,6 m (Fahrstreifenbreite) ausgewählt. Jede der Zellen beinhaltet eine Häufigkeitsverteilung der vorkommenden TTC-Werte bzw. THW-Werte. Diese können Werte in beiden Fällen zwischen 0 s und 20 s mit einer Auflösung von 0,5 s annehmen.

Um das Konzept der Gefahrenerkennung anhand von atypischen Ereignissen geschlossen zu testen, wurden die Objektdaten direkt aus der Verkehrssimulation einbezogen und nicht über die simulierten Sensoren und die Objekterkennung gewonnen. Die Verkehrssimulation beinhaltete Verkehr auf einer dreispurigen geraden Autobahn. Dabei fährt auf dem linken Fahrstreifen schnellerer und auf dem rechten langsamerer Verkehr. Das Ziel bei diesem ersten Experiment war, die Erkennbarkeit von atypischen Situationen in einfachen Szenarien zu prüfen. Daher wurde



**Abbildung 9-13:** Beispiel einer SIG-Struktur. Jede Zelle ist durch den entsprechenden Schwellwert repräsentiert (hell entspricht höheren Werten).



**Abbildung 9-14:** Beispiele für die Verteilungsfunktion in zwei unterschiedlichen SIG-Zellen.

die Verkehrssimulation mit viel Dynamik in der Spurhaltung und der Längsgeschwindigkeit aller Fahrzeuge gestaltet, was im Vergleich zum realen Verkehr eine unnatürliche Situation darstellt.

Die SIG-Strukturen wurden mehrere Stunden lang trainiert mit einem Aktualisierungsintervall von 0,5 s. In Abbildung 9-13 sind die zwei Schichten der resultierenden SIG-Struktur dargestellt. Die Zellen sind nach dem Schwellwert des TTC bzw. THW schattiert. Darin erkennt man, dass nähere Zellen zum Ursprung (Ego-Fahrzeug) einen niedrigeren TTC- bzw. THW-Wert aufweisen (dunkler). Auf der TTC-Schicht ist auch erkenntlich, dass der rechte Verkehr (untere Seite) langsamer fährt als der linke. Auf Abbildung 9-14 sind die Häufigkeitsverteilungen beispielhaft für zwei Zellen für die TTC dargestellt. Darin sind die unterschiedliche Lage und Form der Verteilung in Abhängigkeit von der Entfernung erkennbar. Auch der damit verbundene Schwellwert ist dementsprechend unterschiedlich (TTC von 6,5 s in einer Entfernung von 40 m und 1,5 s in einer Entfernung von 10 m).

Der Prototyp wurde zudem 2 weitere Stunden in der Detektionsphase mit diesen Daten getestet. Dabei ergaben sich 66 Unfälle, 64 davon (96,87%) wurden frühzeitig erkannt. Im Mittel waren die Unfälle 3,2 s vorher als Gefahr detektiert. Die Rate an Warnungen war mit 161 pro Stunde sehr hoch. Dies ist auf die bereits erwähnten unnatürlichen Verhältnisse der Verkehrssimulation zurückzuführen.

Mit diesen Experimenten ist der Ansatz der Erkennung von atypischen Situationen in einfachen Szenarien getestet worden. Eine Validierung der Ergebnisse mit realen Daten und Fahrereinschätzungen wurden im Rahmen dieser Arbeit nicht vorgenommen.



## 10 Zusammenfassung und Ausblick

### 10.1 Zusammenfassung und Diskussion

In der vorliegenden Arbeit wurde eine Prozessierungskette zur Objekterkennung und Verfolgung in Fahrsituationen sowie zur Ableitung von potentiellen Gefahren vorgestellt. Im Zusammenhang mit dieser Prozessierungskette sind an vielen Stellen neue Methoden vorgestellt worden, einschließlich der Verarbeitung auf Sensorebene, der Objekterkennung aus Kamerabildern und Laserscannerdaten, der Objektverfolgung und Fusion sowie der Gefahrenerkennung.

In Anlehnung an die in Abschnitt 1.5 gestellten Forschungsfragen werden hier die erarbeiteten Punkte wieder aufgegriffen und diskutiert. Dabei lautet die zentrale Frage: *welche Chancen und Probleme stellt die Fusion der Daten einer Kamera und eines Laserscanners in der Detektion von Objekten im unmittelbaren Umfeld des Fahrzeugs im Blick auf die frühzeitige Erkennung von potentiellen Gefahren in der Fahrsituation dar?* Bereits in den einleitenden Kapiteln der Arbeit wurde erkannt und herausgestellt, dass die erwähnten Sensoren viele Stärken zur Objekt- und Gefahrenerkennung besitzen. Da die Sensoren auf unterschiedlichen Prinzipien beruhen, kann durch Fusion ihrer Daten ein größerer Gewinn für die Objekterkennung, und damit auch für die Gefahrenerkennung, erwartet werden. Allerdings sind ebenso wichtige Schwierigkeiten erkannt worden, die in den folgenden Absätzen zusammengefasst werden.

#### 10.1.1 Objektverfolgung und Fusion

*Wie können die Daten aus einem oder unterschiedlichen Sensoren für die Objekterkennung und Verfolgung so verarbeitet bzw. fusioniert werden, dass qualitativ hochwertige Objekthypothesen entstehen? Diese sollen zudem robust sein, gegenüber*

- *sich verändernden Perspektiven der Objekte,*
- *manövrierenden Objekten*
- *und unsicheren aktuellen Sensordaten.*

Im Zusammenhang mit dieser Frage ist der Kern der Arbeit gewachsen. Die grundlegende Struktur entspricht dabei einem Multi-Target-Tracking-Schema auf Objektebene, wobei aus jedem Sensor Objektbeobachtungen geliefert werden. Durch die Fusion auf Objektebene werden eine hohe Modularität, Verteilbarkeit und Wiederverwendbarkeit gewährleistet. Insbesondere wird durch die Abkopplung von den Sensoren eine hohe Systemverfügbarkeit erreicht, da sowohl mit einem als auch mit beiden (oder mehreren) Sensoren gearbeitet werden kann. Hierzu wurden ein einfaches Quadermodell für Objekthypothesen und Beobachtungen sowie ein geradliniges Bewegungsmodell mit konstanter Geschwindigkeit ausgewählt. Dadurch kann die Objektverfolgung (anhand eines Informationsfilters) schnell initialisiert und stabil durchgeführt werden.

Um die hohen Ansprüche der Gefahrenerkennung im Bezug auf die in der Frage aufgelisteten Aspekte zu gewährleisten, sind folgende zusätzliche Strategien und entsprechende Methoden

entwickelt worden. Um Objekte bei sich verändernder Perspektive gut erkennen und verfolgen zu können, wurde eine variable, logische Referenz der Objekte und partielle Beobachtbarkeit eingeführt. Dabei werden Objekte sehr detailliert beschrieben und kompakt dargestellt. Des Weiteren wurden zur Verfolgung von manövrierenden Objekten (nicht erfassbar durch das einfache Bewegungsmodell) neben der Anpassung des Prozessrauschens ein Multi-Assoziationsverfahren und eine Verdoppelungsstrategie entwickelt. Schließlich wurde zum Umgang mit Ausreißern und *split*- und *merge*-Effekten eine Kombination aus den bisher aufgeführten Methoden mit einem Vereinigungsmechanismus entwickelt.

Durch diese Strategien und Methoden konnten in dieser Arbeit die in der oberen Frage aufgelisteten Aspekte deutlich verringert werden. Durch Fusion der Daten mehrerer Sensoren wird trotz erhöhter Komplexität ein Gewinn in der Robustheit, Verfügbarkeit und Zuverlässigkeit der Objektverfolgung erreicht. Allerdings zeigt sich dabei eine zum Teil geringfügig niedrigere Genauigkeit der fusionierten Daten. In der Optimierung der Genauigkeit in Kombination mit den Vorteilen der hier vorgestellten Methoden liegt weiteres Forschungspotential für zukünftige Aktivitäten.

### 10.1.2 Objekterkennung

*Wie kann die notwendige Verarbeitung der Kamera- und Laserscannerdaten die Anforderungen der Objekterkennung weitestgehend erfüllen? Wie kann die Menge der zu verarbeitenden Sensordaten auf den wesentlichen Bereich reduziert werden?*

Um die soeben vorgestellte Objektverfolgung und Fusion auf Objektebene zu unterstützen, ist es wichtig, sichere, genaue und wesentliche Objektbeobachtungen aus den Sensordaten zu extrahieren. Hierzu sind Verarbeitungsmethoden für jeweils Kamerabilder und Laserscannerdaten entwickelt und vorgestellt worden. Zur Verarbeitung der Kamerabilder wurde ein möglichst generischer und modellfreier Ansatz entwickelt, nämlich die Bildsegmentierung anhand eines Texturmerkmals. Der Vorteil dieser Methode gegenüber anderen Arbeiten ist, dass sie keine anderen Annahmen über die Objekte trifft, außer dass sie sich vom Hintergrund abheben. Vor der Objektextraktion werden die Fahrzeugschwankungen im Bild anhand einer neuartigen Bildverarbeitungsmethode korrigiert, die auf der Geometrie der Fahrstreifenmarkierungen basiert.

Zur Verarbeitung der Laserscannerdaten wurde eine Methode eingeführt, mit der die Objektkonturen schnell und differenziert an das Quadermodell angepasst werden. Zusätzlich wurden besondere Algorithmen entworfen, um die logische Referenz der Objekte, die partielle Beobachtbarkeit sowie die angepasste Messunsicherheit zu bestimmen.

Um die Aufmerksamkeit dieser Verarbeitung auf den Straßenbereich zu fokussieren, wurde ebenso eine neuartige Methode zur Ermittlung der Fahrkorridor Grenzen entwickelt. Dank einer Fusionierung hat sie eine hohe Zuverlässigkeit und trotzdem eine hohe Performanz. Mit dieser Methode werden der Rechenaufwand der Objekterkennung reduziert und zahlreiche Rauschobjekte in den seitlichen Bereichen unterdrückt.

Mit dieser Menge an neuen Methoden konnten die oberen Fragen weitestgehend beantwortet werden. Es wurden Objekterkennungsmethoden aus den jeweiligen Sensordaten entwickelt, die eine hohe Zuverlässigkeit und Genauigkeit besitzen und dabei an die im letzten Abschnitt vorge-



stellten Strategien zur Objektverfolgung angepasst wurden. Auch die Menge an Daten konnte dank der Fahrkorridorerkennung auf den wesentlichen Bereich eingeschränkt werden.

### 10.1.3 Gefahrenerkennung

*Wie müsste die Architektur einer allgemeingültigen aber gleichzeitig differenzierten Gefahrenerkennung aufgebaut sein, die möglichst frühzeitig auf entsprechende Ereignisse reagiert? Wie hängt die Güte der Objektdaten mit der der Gefahrenerkennung zusammen?*

Das letztliche Ziel der Sensordatenverarbeitung und Fusion besteht in der Verwendung der gewonnenen Daten auf Anwendungsebene, wie beispielsweise zur Erkennung von potentiellen Gefahrensituationen. Es wurde in dieser Arbeit ein statistischer Ansatz zur Ableitung von potentiellen Gefahren in der Fahrsituation entwickelt. Darin werden Gefahren aufgrund ihrer Seltenheit erkannt. Zur Behandlung der statistischen Informationen (SIG) muss jedoch ein Kompromiss zwischen der Dimension der Speicherstrukturen (so gering wie möglich) und der Komplexität der erkennbaren Gefahren (so hoch wie möglich) getroffen werden. Es wurde in dieser Arbeit vorgeschlagen, jede Komponente der Fahrsituation getrennt zu betrachten, womit keine Korrelationen berücksichtigt und keine hoch komplexen Gefahren erkannt werden können. Dabei bleibt aber der Ansatz schlank und abgekoppelt von den Eingangsdaten, was deutlich allgemeiner und flexibler ist. Außerdem ist aufgrund der Differenzierung im SIG eine frühzeitige Erkennung von potentiellen Gefahren möglich.

Die praktische Umsetzung des Ansatzes wurde anhand der Objektdaten (als TTC oder THW) diskutiert. Dabei wurde besonderer Wert auf den Zusammenhang zwischen der Güte der eingehenden Objektdaten und der Güte der resultierenden Gefahrenerkennung gelegt. Es konnte der direkte Einfluss des Erfassungsbereichs, der Genauigkeit, der Verfügbarkeit und der Zuverlässigkeit auf die resultierenden Schwellwerte der Gefahrenerkennung belegt werden.

Mit dem vorgestellten Ansatz konnten auf die oberen Fragen Antworten gegeben werden, insbesondere im Blick auf die allgemeingültige, differenzierte und frühzeitige Erkennung von potentiellen Gefahrensituationen. Die Komplexität der erkennbaren Gefahren im Bezug auf diesen Ansatz bleibt ein interessanter Aspekt für zukünftigen Fragestellungen.

## 10.2 Ausblick

Mit dieser Arbeit ist ein Beispiel für eine Verarbeitungskette von der Datenerfassung, über Verarbeitungsmethoden (Erkennung und Fusion von Objekten), bis hin zur Anwendungsebene (Gefahrenerkennung) vorgestellt worden. Aufgrund der Breite dieses Feldes konnten lediglich die wesentlichen Problemstellungen behandelt werden. Es sind dabei viele neue Fragestellungen aufgeworfen worden, die jedoch im zeitlichen Rahmen dieser Arbeit nicht umfassend abgehandelt werden konnten. Diese stellen interessante Herausforderungen für weitere Arbeiten im vorgestellten Themenbereich dar, welche im Folgenden zusammengefasst werden.

Es wäre aufbauend auf dieser Arbeit sinnvoll, die vorgestellten Methoden in allen Ebenen mit weiteren Merkmalen, Testszenarien, Bedingungen und anderen Sensoren oder Sensorkombinationen zu festigen. Beispielsweise könnten mit einem Radar oder einer Stereo-Vision die Stärken

des Systems ergänzt werden. Weitere Möglichkeiten stellen infrastrukturseitige Sensoren sowie kommunikationsbasierte Informationen (*Car2X*) dar. Auch das Testen des Systems auf Landstraßen und in innenstädtischen Bereichen könnte wichtige Informationen über die Methoden liefern und neue Herausforderungen durch die höhere Anzahl an Fahrzeugen und die komplexere Struktur stellen.

Weitere Fragen stellten sich im Laufe der Arbeit. Diese beziehen sich beispielsweise auf die Beibehaltung der Kalibrationseigenschaften der Sensoren, auf die in Kapitel 9 festgestellten Unterschiede zwischen den Objektbeobachtungen der Sensoren sowie auf die tiefgreifende Charakterisierung des (weißen) Mess- und Prozessrauschens.

Auch die Gefahrenerkennung könnte weiter untersucht werden. Hierzu wäre ein Abgleich der erkannten Gefahren mit der Fahrerwahrnehmung sowie mit Daten aus Unfalldatenbanken sinnvoll. Insbesondere in einem Kreuzungsbereich besitzt die Gefahrenerkennung viel Potential. Dazu sollte getestet werden, wie sich die detektierten und verfolgten Objektdaten für die Aufgabe der Gefahrenerkennung eignen.

Zukünftig werden zahlreiche der hier entwickelten Methoden und angesprochenen Ideen in weiteren Projekten am Institut TS umgesetzt. Insbesondere wird die Integration der Methoden in das Versuchsfahrzeug *FASCar* weiterentwickelt. Hierzu werden viele der hier aufgelisteten offenen Punkte und Fragen zu betrachten sein. Es bleibt von großem Interesse, welche weiteren Fortschritte in Zukunft auf der Basis dieser Arbeit erreicht werden können.

## A Sensormodelle

In diesem Anhang werden Grundlagen und geometrische Modelle über die in dieser Arbeit verwendeten Sensoren, die Kamera und der Laserscanner vorgestellt.

### A.1 Sensorik und Orientierungsparameter

Sensoren bestehen üblicherweise aus einer Anordnung von Detektoren und anderen Elementen (wie Motoren, Linsen, Filter etc.). Beispielhaft sind in Abbildung A-1 vier Sensoren mit ihrem schematischen Aufbau dargestellt. Eine Kamera besitzt einen Chip mit einer Anzahl von Pixeln (lichtempfindlichen Detektoren), sowie eine Blende und ein Objektiv (Linsensystem). Ein Laserscanner besteht aus einem laserbasierten Abstandsmesser (Lidar) mit einem davor montierten rotierenden Spiegel und liefert die reflektierte Strahlung (Entfernung) zu Objekten in jeder Richtung. Einige Modelle von Laserscannern nutzen mehrere Laserstrahlen, womit gleichzeitig mehrere Ebenen gescannt werden können.

Jeder Sensor besitzt eine Empfindlichkeit, die sich in Form eines Wertbereichs (Min- und Max-Werte) und einer Granularität (feinste Erkennungsstufe) ausdrücken lässt. Daraus ergibt sich die Quantisierung der Sensordaten. Die von einem Sensor gelieferten Daten sind auch durch die zeitliche und räumliche Abtastung (*Sampling*) bestimmt. So liefert eine Kamera eine bestimmte Anzahl von Bildern pro Sekunde in einer bestimmten geometrischen, spektralen und radiometrischen Auflösung. Diese Parameter lassen sich zum Teil im Sensor einstellen oder durch Kalibrierung des Sensors feststellen.

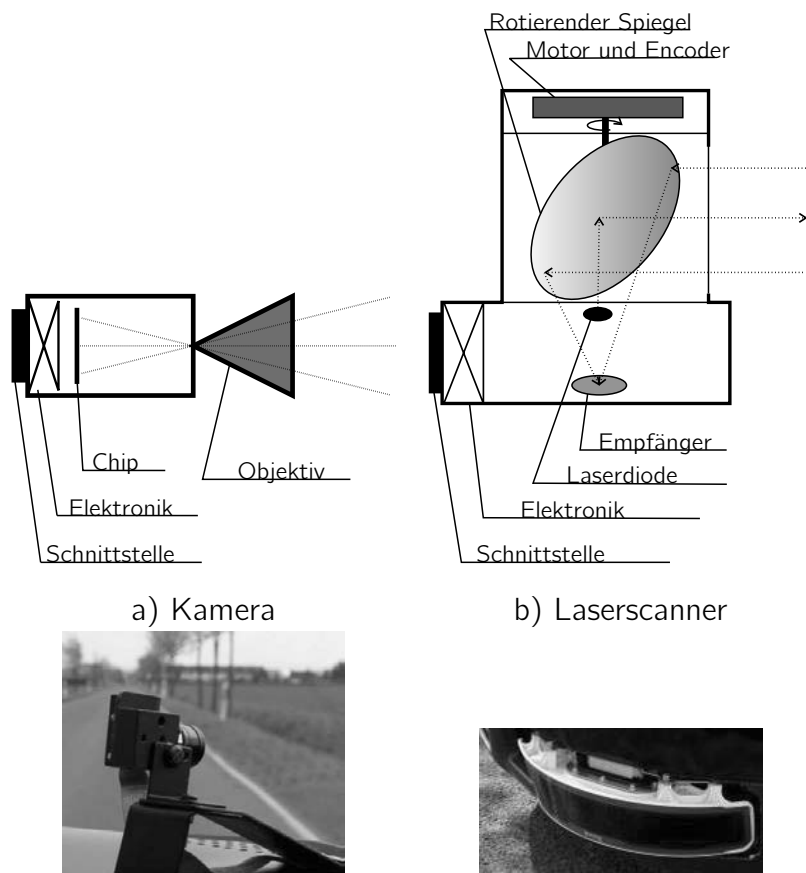
Durch die Anordnung der Elemente eines komplexen Sensors (innere Orientierung) ist sein lokaler Erfassungsbereich bestimmt. Für eine Kamera sind diese Parameter beispielsweise die Brennweite, der Bildhauptpunkt, die Pixelgröße, das Seitenverhältnis und die Linsenparameter (siehe Details in Abschnitt A.2). Für einen Laserscanner sind die Parameter unter anderen die Hauptachse und der Start- sowie Enderfassungswinkel.

Durch die Anordnung des Sensors zur Welt, durch feste oder bewegliche Montierung, ist der Bezug des lokalen zum globalen Erfassungsbereich, also die äußere Orientierung, gegeben. Diese besteht aus der Position und der Ausrichtung des Sensors. In der äußeren Orientierung kann auch die Trägerplattform als Zwischentransformation betrachtet werden (Alignment-Parameter<sup>1</sup>). Abbildung A-2 verdeutlicht den Zusammenhang zwischen dem Fahrzeugkoordinatensystem  $[x_v \ y_v \ z_v]^T$  (mit Ursprung in der Mitte der Fahrzeugsvorderachse) und dem Sensorkoordinatensystem, in diesem Fall einer Kamera  $[x_c \ y_c \ z_c]^T$ , anhand der Translation  $(t_x, t_y, t_z)$  und der Rotationswinkel (um die  $Z$ -Achse  $\kappa$ , um die  $Y$ -Achse  $\phi$  und um die  $X$ -Achse  $\omega$ , in dieser Reihenfolge, mit positivem Winkel in Uhrzeigersinn). Diese Beschreibung gibt die zugrunde liegende Konvention für diese Arbeit wieder.

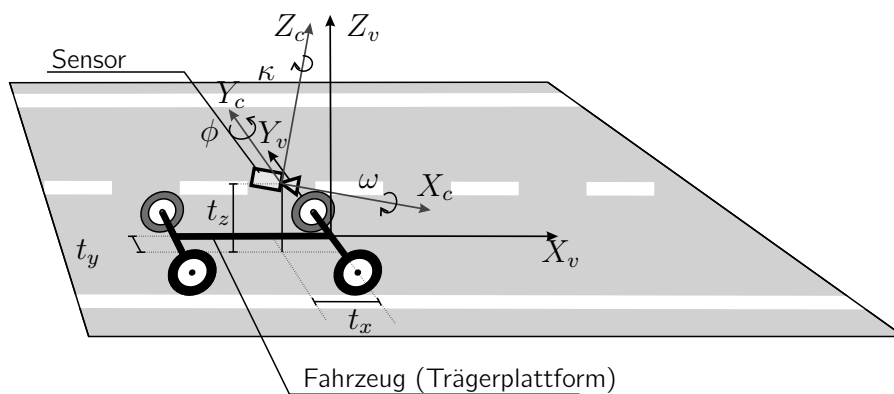
Die relative Orientierung der Sensoren auf der Trägerplattform muss bekannt sein, um deren

---

<sup>1</sup>Die Alignment-Parameter sind meist fest, auch wenn sich Einbauort sowie Ausrichtung beispielsweise durch Erschütterungen ungewollt verändern können.



**Abbildung A-1:** Kamera und Laserscanner Sensoren mit schematischer Aufbau.



**Abbildung A-2:** Abbildungsvorgang in einem Sensor unter Berücksichtigung der Alignment-Parameter.

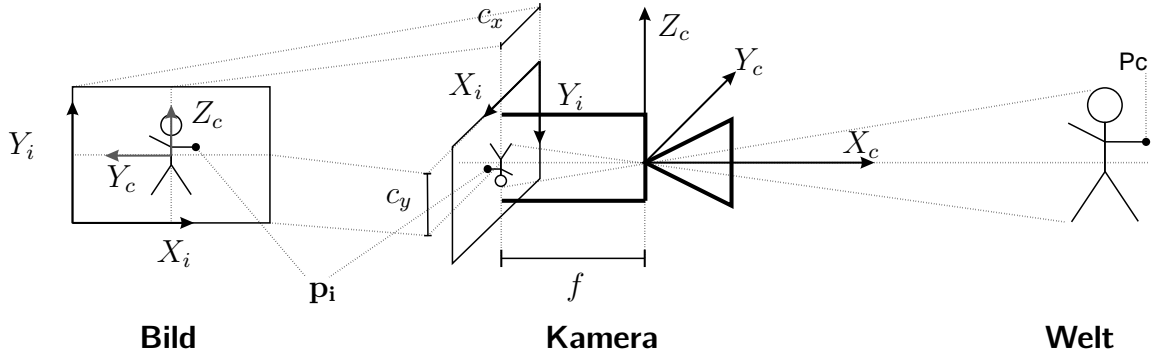


Abbildung A-3: Abbildungsvorgang in einer idealen Lochkamera.

Daten in einem gemeinsamen Referenzraum darstellen zu können.

## A.2 Geometrische Beschreibung einer Kamera

In der Abbildung A-3 ist der Abbildungsvorgang einer idealisierten Lochkamera anhand der inneren Orientierung vorgestellt, so dass das Weltkoordinatensystem auf der Kamera zentriert ist, also Kamerakoordinaten (Index  $c$  wie *camera*). Der Ursprung ist im Projektionszentrum (Kameraloche) und die Achsen sind, ohne Verlust an Allgemeinheit, wie üblicherweise im Fahrzeug,  $Z$  nach oben,  $Y$  nach links und  $X$  nach vorn ausgewählt. Auf der linken Seite ist die Abbildung in Bildkoordinaten dargestellt. Ist ein Punkt in der Szene in homogenen Koordinaten<sup>2</sup> als  $\mathbf{p}_c = [x_c \ y_c \ z_c \ 1]^T$  gegeben, so wird dieser anhand folgender Gleichung in einen Bildpunkt  $\mathbf{p}_i = [x_i \ y_i \ 1]^T$  abgebildet (Index  $i$  wie *image*). In dieser Arbeit sind Weltkoordinaten immer in Millimeter und Bildkoordinaten in Pixel gegeben.

$$\mathbf{p}_i = \mathbf{T}_c^i \cdot \mathbf{p}_c \quad (\text{A.1})$$

$$\begin{bmatrix} x'_i \\ y'_i \\ q \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \quad (\text{A.2})$$

Hier sind  $x_i = x'_i/q$  und  $y_i = y'_i/q$  die resultierenden Bildkoordinaten. Darin bezeichnet  $(f_x, f_y)$  die Brennweite des Objektivs<sup>3</sup> und  $(c_x, c_y)$  der Bildhauptpunkt. Die Brennweite ist in der Abbildung A-3 als  $f$  dargestellt. Um normalisierte Koordinaten zu erzeugen, muss das Seitenverhältnis  $\alpha$  (*aspect ratio*, Breite zu Höhe, z.B. 4:3) und die Pixelgröße  $\mu$  (in Millimeter) für beide Komponenten mitberücksichtigt werden. Da in dieser Arbeit die Bildkoordinaten  $X$  nach rechts und  $Y$  nach oben zeigen, muss dementsprechend die Richtung der horizontalen Achse umgekehrt werden (durch  $-1$  multipliziert), wie in der Abbildung A-3 links verdeutlicht. Die Ausdrücke sind  $f_x = -\alpha f/\mu$  und  $f_y = f/\mu$ . Um den Bildkoordinatenursprung auf eine der Bildecken (z.B. unten links) zu versetzen, werden die Bildkoordinaten des Hauptpunkts verwendet.

<sup>2</sup>In homogenen Koordinaten entspricht  $[x \ y \ z \ q]^T = [x/q \ y/q \ z/q \ 1]^T$

<sup>3</sup>In einem photogrammetrischen System wird statt Brennweite der Begriff Kamerakonstante verwendet, womit der Abstand zwischen Projektionszentrum und Bildebene gemeint ist. Die beiden betragen in der Regel annähernd gleiche Werte, wenn das Objektiv auf Unendlich eingestellt ist.

Da die Kamera im Fahrzeug eingebaut ist, ist es wünschenswert, deren Koordinaten zu den Fahrzeugkoordinaten (Plattformkoordinaten, Index  $v$  wie *vehicle*) anhand der Alignment-Parameter (äußere Orientierung) zu transformieren,  $\mathbf{p}_v = [x_v \ y_v \ z_v \ 1]^T$ . Dieses Koordinatensystem, siehe Abbildung A-3 auf der vorherigen Seite, ist analog zum Kamerasystem, mit  $Z$  nach oben,  $Y$  nach links und  $X$  nach vorn ausgewählt. Dessen Ursprung liegt in der Mitte der vorderen Radachse. Für die Transformation müssen die Alignment-Parameter bekannt sein, in diesem Fall die drei Rotationswinkel und die Translationsparameter vom Ursprung zum Projektionszentrum ( $t_x$ ,  $t_y$  und  $t_z$ ). Diese affine Transformation wird in homogenen Koordinaten wie folgt ausgedrückt

$$\mathbf{p}_c = \mathbf{T}_v^c \cdot \mathbf{p}_v \quad (\text{A.3})$$

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{T}_{\text{rot}} & \mathbf{T}_{\text{trans}} \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_v \\ y_v \\ z_v \\ 1 \end{bmatrix}, \quad (\text{A.4})$$

wobei  $\mathbf{T}_{\text{trans}}$  eine  $3 \times 1$  Translationsmatrix mit der Form  $\mathbf{T}_{\text{trans}} = [-t_x \ -t_y \ -t_z]^T$  und  $\mathbf{T}_{\text{rot}}$  eine  $3 \times 3$  Rotationsmatrix entsprechend eines Rechtssystems

$$\mathbf{T}_{\text{rot}} = \mathbf{T}_{\text{rotX}} \cdot \mathbf{T}_{\text{rotY}} \cdot \mathbf{T}_{\text{rotZ}} \quad (\text{A.5})$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & c\omega & s\omega \\ 0 & -s\omega & c\omega \end{bmatrix} \begin{bmatrix} c\phi & 0 & s\phi \\ 0 & 1 & 0 \\ -s\phi & 0 & c\phi \end{bmatrix} \begin{bmatrix} c\kappa & s\kappa & 0 \\ -s\kappa & c\kappa & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (\text{A.6})$$

mit  $c\omega = \cos(\omega)$ ,  $s\omega = \sin(\omega)$ ,  $c\phi = \cos(\phi)$ ,  $s\phi = \sin(\phi)$ ,  $c\kappa = \cos(\kappa)$ , und  $s\kappa = \sin(\kappa)$  ist.

Analog dazu kann die Koordinatentransformation anhand der äußeren Orientierung von Fahrzeug in Weltkoordinaten (beispielsweise UTM-System<sup>4</sup>, Index  $w$  wie *world*) umgerechnet werden,  $\mathbf{T}_w^v$ .

Und somit kann die gesamte Transformation zwischen Welt- und Bildkoordinaten als Matrixmultiplikation wie folgt dargestellt werden:

$$\mathbf{p}_i = \mathbf{T}_c^i \cdot \mathbf{T}_v^c \cdot \mathbf{T}_w^v \cdot \mathbf{p}_w \quad (\text{A.7})$$

Diese Ausdrücke sind für eine ideale Lochkamera ausgelegt. Bei realen Kameras (insbesondere bei Objektiven mit kurzen Brennweiten) treten jedoch Linsenverzerrungen auf, womit die innere Orientierung, Gleichung A.1, angepasst werden muss. Unter den Verzerrungsmodellen zählt das von Brown [22] zu den meist verwendeten (radiale und tangential Verzerrungen).

<sup>4</sup>UTM-Koordinatensystem (*Universal Transverse Mercator*)

### A.3 Geometrische Beschreibung eines Laserscanners

Üblicherweise sind die inneren Parameter eines Laserscanners vom Hersteller bestimmt worden, so dass die Ausgangsdaten des Laserscanners bereits in sensorzentrierten kartesischen Koordinaten ausgegeben werden. Dabei sind unter anderen die Strahlenlenkung, sowie die richtungsabhängige Winkelauflösung berücksichtigt.

In einem vereinfachten Modell können die Rohdaten eines Laserscanners in Polarkoordinaten durch die Entfernung  $\rho$  eines Reflexionspunktes aus einer bestimmten Richtung  $\theta$  gegeben sein. In diesem Fall entspricht die nichtlineare Transformation in kartesischen Koordinaten (I für *laser scanner*):

$$\mathbf{p}'_I = \begin{bmatrix} \rho & \theta \end{bmatrix}^T \leftrightarrow \mathbf{p}_I = \begin{bmatrix} x & y \end{bmatrix}^T \quad (\text{A.8})$$

$$\mathbf{p}'_I = \begin{bmatrix} \sqrt{x^2 + y^2} \\ \arctan \frac{y}{x} + \theta_c \end{bmatrix}^T \quad \mathbf{p}_I = \begin{bmatrix} \rho \cdot \cos(\theta + \theta_c) \\ \rho \cdot \sin(\theta + \theta_c) \end{bmatrix}^T \quad (\text{A.9})$$

Der Hauptparameter ist die Nullausrichtung des Laserscanners, hier durch  $\theta_c$  verdeutlicht. Bei der Transformation in Polarkoordinaten ist zu beachten, dass der Winkel  $\theta$  im Bereich  $(-\pi, \pi]$  sein muss. Die Grenzen der Messungen sind durch den Winkelbereich  $[\theta_{min}, \theta_{max}]$  und Entfernungsbereich  $[\rho_{min}, \rho_{max}]$  gegeben.

Dieses Modell ist vereinfacht. Betrachtet man den Aufbau eines Laserscanners (siehe Abbildung A-1 auf Seite 132), so sollten weitere Parameter berücksichtigt werden, wie zum Beispiel der Abstand zwischen Sender und Spiegel, der Winkel des Spiegels (idealerweise  $45^\circ$ ), die Verschiebung der Senderposition bezüglich der Spiegel-Drehachse (idealerweise 0), die Abweichung zwischen Sendewinkel und Drehachse, und Unebenheiten des Spiegels. Außerdem sollen die unterschiedlichen Scanebenen gesondert modelliert werden. Die meisten Hersteller verarbeiten die Daten so, dass die Daten in kartesischen Koordinaten ( $\mathbf{p}_I$  oder unter Berücksichtigung der Alignment-Parameter  $\mathbf{p}_w$ ) herausgegeben werden. Gordon beschäftigt sich in ihrer Dissertation [66] mit dem Einfluss dieser und anderer Faktoren.

Des Weiteren können die Laserscannerkoordinaten in Weltkoordinaten transformiert werden (analog zur Gleichung A.7):

$$\mathbf{p}_I = \mathbf{T}_v^I \cdot \mathbf{T}_w^v \cdot \mathbf{p}_w \quad (\text{A.10})$$





## B Geometrische Transformation zwischen der Bild- und der Stra enebene

Die in einer idealen Lochkamera stattfindende Transformation zwischen einem Punkt in Fahrzeugkoordinaten  $\mathbf{p}_v$  (*vehicle*) und Bildkoordinaten  $\mathbf{p}_i$  (*image*) kann in homogenen Koordinaten als eine nicht invertierbare  $4 \times 3$  Matrix  $\mathbf{T}_v^i$  ausgedr ckt werden (siehe letzten Abschnitt):

$$\mathbf{p}_i = \mathbf{T}_v^i \cdot \mathbf{p}_v \quad (\text{B.1})$$

$$\mathbf{T}_v^i = \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{24} \\ t_{31} & t_{32} & t_{33} & t_{34} \end{bmatrix} \quad (\text{B.2})$$

Die Stra enkoordinaten entsprechen den Fahrzeugkoordinaten auf einer bestimmten H he, normalerweise  $z_v = 0$ , und werden hier mit Index  $r$  wie *road* dargestellt. Betrachtet man die Stra enebene, so kann die  $Z$ -Komponente aus  $\mathbf{p}_v$  als Konstante in die Matrix eingebunden werden. Somit erh lt man eine quadratische und nichtsingul re Matrix  $\mathbf{T}_r^i$ , die invertierbar ist (Transformation zwischen zwei Ebenen). Die zweidimensionalen Stra enpunkte werden als  $\mathbf{p}_r$  bezeichnet:

$$\mathbf{p}_i = \mathbf{T}_r^i \cdot \mathbf{p}_r \Leftrightarrow \mathbf{p}_r = \mathbf{T}_r^{i^{-1}} \cdot \mathbf{p}_i = \mathbf{T}_i^r \cdot \mathbf{p}_i \quad (\text{B.3})$$

$$\mathbf{T}_r^i = \begin{bmatrix} t_{11} & t_{12} & t_{13} \cdot z_v + t_{14} \\ t_{21} & t_{22} & t_{23} \cdot z_v + t_{24} \\ t_{31} & t_{32} & t_{33} \cdot z_v + t_{34} \end{bmatrix} \quad (\text{B.4})$$

Diese Transformation kann ebenso f r Geraden angepasst werden. Sei eine Bildgerade  $\mathbf{l}_i = [A_i \ B_i \ C_i]^T$  in homogenen Bildkoordinaten gegeben ( $A_i x + B_i y + C_i = 0$ ), so dass f r die beinhalteten Punkte  $\mathbf{p}_i$  gilt:

$$\mathbf{l}_i^T \cdot \mathbf{p}_i = 0 \Leftrightarrow \mathbf{l}_i^T \cdot \mathbf{T}_r^i \cdot \mathbf{p}_r = 0 \quad (\text{B.5})$$

So wird die Gerade in der Stra enebene zu

$$\mathbf{l}_r^T = \mathbf{l}_i^T \cdot \mathbf{T}_r^i \leftrightarrow \mathbf{l}_r = \mathbf{T}_r^{iT} \cdot \mathbf{l}_i = \mathbf{T}_i^r \cdot \mathbf{l}_i. \quad (\text{B.6})$$

Dabei ist f ur die Geradentransformation von Stra enebene auf Bildebene keine Matrixinversion notwendig.

## C Der Kalman-Filter und der Informationsfilter

In diesem Anhang werden die Herleitungen des Kalman-Filters und des Informationsfilters gegeben (Quellen: Brown and Hwang [23], Bar-Shalom [10] und Mutambara [109]).

### C.1 Systembeschreibung

Zur Aufgabe der Umfelderkennung werden Zustandsinformationen anhand von indirekten, unsicheren und ungenauen Messdaten bzw. Beobachtungen aus den Sensoren oder der Verarbeitung geschätzt. Hier wird der Fokus auf lineare dynamische zeitdiskrete Systeme gelegt, mit

$$t_k = t_{k-1} + \Delta t \quad (\text{C.1})$$

Die zu schätzende Information wird in jedem Zeitpunkt  $t_k$  als Systemzustand  $\mathbf{x}(t_k)$  mit  $n_x$  Komponenten dargestellt:

$$\mathbf{x}(t_k) = \begin{bmatrix} x_1 & x_2 & \cdots & x_{n_x} \end{bmatrix}^T \quad (\text{C.2})$$

Der Schätzfehler entspricht dabei

$$\mathbf{e}(t_k) = \tilde{\mathbf{x}}(t_k) - \hat{\mathbf{x}}(t_k), \quad (\text{C.3})$$

mit  $\tilde{\mathbf{x}}$  dem wahren Wert und  $\hat{\mathbf{x}}$  der Schätzung des Zustands. Die entsprechende Fehlerkovarianzmatrix, unter Annahme einer erwartungstreuen Schätzung, lautet

$$\mathbf{P}(t_k) = \mathbb{E}[\mathbf{e}(t_k)\mathbf{e}^T(t_k)] \quad (\text{C.4})$$

und wird mitgeschätzt. Das Systemverhalten lässt sich als Zufallsprozess im Zeitbereich durch ein funktionales Modell beschreiben:

$$\mathbf{x}(t_k) = \mathbf{f}[t_k, \mathbf{x}(t_{k-1}), \mathbf{u}(t_k), \mathbf{w}(t_k)] \quad (\text{C.5})$$

Ist dieses linear, so wird es in Matrixform wie folgt ausgedrückt:

$$\mathbf{x}(t_k) = \mathbf{F}(t_k) \cdot \mathbf{x}(t_{k-1}) + \mathbf{B}(t_k) \cdot \mathbf{u}(t_k) + \mathbf{w}(t_k) \quad (\text{C.6})$$

Dabei stellen  $\mathbf{u}(t_k)$  den Steuereingang des Systems und  $\mathbf{B}(t_k)$  die Steuerungsmatrix dar.  $\mathbf{F}(t_k)$  ist die Systemmatrix.  $\mathbf{w}(t_k)$  stellt das Prozessrauschen (stochastisches Modell) des Systems dar. Dieses wird als zeitlich unkorrelierter, mittelwertfreier, normalverteilter Zufallsprozess mit Kovarianzmatrix  $\mathbf{Q}(t_k)$  modelliert (weißes Rauschen<sup>1</sup>):

$$\mathbb{E}[\mathbf{w}(t_k)] = \mathbf{0}; \mathbb{E}[\mathbf{w}(t_k)\mathbf{w}(t_i)] = \begin{cases} \mathbf{Q}(t_k) & \text{wenn } k = i \\ 0 & \text{sonst.} \end{cases} \quad (\text{C.7})$$

Die Sensoren liefern in jedem Zyklus  $N_z$  Messwerte (auch Beobachtungen genannt) in einem Messraum. Ein Messvektor des Sensors  $s$  wird als  $\mathbf{z}_s(t_k)$  mit  $n_z$  Komponenten ausgedrückt:

$$\mathbf{z}_s(t_k) = \begin{bmatrix} z_1 & z_2 & \cdots & z_{n_z} \end{bmatrix}^T \quad (\text{C.8})$$

Der Zusammenhang zwischen dem Messvektor des Sensors  $s$  und dem Systemzustand ist durch die Funktion  $\mathbf{h}_s[t_k, \mathbf{x}(t_k), \mathbf{v}(t_k)]$  gegeben. Für viele Systeme ist diese linear und kann in Form einer Beobachtungsmatrix  $\mathbf{H}_s$ ,  $n_z \times n_x$ , ausgedrückt werden.

$$\mathbf{z}_s(t_k) = \mathbf{H}_s(t_k) \cdot \mathbf{x}(t_k) + \mathbf{v}_s(t_k) \quad (\text{C.9})$$

Dabei stellt  $\mathbf{v}_s(t_k)$  das Messrauschen des Sensors dar. Dieses wird analog zu  $\mathbf{w}(t_k)$  als weißes Rauschen mit Kovarianzmatrix  $\mathbf{R}_s(t_k)$  modelliert. Dazu müssen die systematischen Fehler bereits korrigiert worden sein. Oft ist die Beobachtungsmatrix nicht invertierbar, so dass sich der Systemzustand aus der Beobachtung nicht direkt errechnen lässt.

## C.2 Der Kalman-Filter

Basierend auf dieser Systembeschreibung wird hier der Kalman-Filter [78] als rekursive Lösung der diskreten linearen Filterung eingeführt. Die folgende Herleitung wird an Brown und Hwang [23] angelehnt.

Der zeitdiskrete Kalman-Filter besteht aus zwei Schritten. Zuerst wird der Zustand  $\hat{\mathbf{x}}$  auf den aktuellen Zeitpunkt prädiziert. Im zweiten Filterschritt, der Schätzung, wird die Innovation berechnet, die durch die Beobachtung in den Filter eingeht. Zur Gewichtung der Innovation wird

---

<sup>1</sup>Weißes Rauschen: Zufallsprozess, dessen Autokovarianz für zwei beliebige unterschiedliche Zeitpunkte null ist.

die Kalman-Matrix  $\mathbf{K}$  (auch *Gain*-Matrix) verwendet.

Zur Prädiktion wird das Dynamikmodell (C.6) (hier ohne Steueranteile) verwendet:

$$\hat{\mathbf{x}}(t_k|t_{k-1}) = \mathbf{F}(t_k)\hat{\mathbf{x}}(t_{k-1}|t_{k-1}) \quad (\text{C.10})$$

Dabei wird auf der Annahme beruht, dass das Prozessrauschen  $\mathbf{w}$  mittelwertfrei und zeitlich unkorreliert ist. Die Bedingung  $t_k|t_{k-1}$  verdeutlicht, dass nur die Beobachtungen bis zum Zeitpunkt  $t_k$  in Form der letzten Schätzung in den Filter eingehen.

Der prädizierte Fehler entspricht:

$$\begin{aligned} \mathbf{e}(t_k|t_{k-1}) &= \tilde{\mathbf{x}}(t_k) - \hat{\mathbf{x}}(t_k|t_{k-1}) \\ &= \mathbf{F}(t_k)\tilde{\mathbf{x}}(t_{k-1}) + \mathbf{w}(t_k) - \mathbf{F}(t_k)\hat{\mathbf{x}}(t_{k-1}|t_{k-1}) \\ &= \mathbf{F}(t_k)\mathbf{e}(t_{k-1}) + \mathbf{w}(t_k) \end{aligned} \quad (\text{C.11})$$

Damit lautet die entsprechende prädizierte Fehlerkovarianzmatrix:

$$\begin{aligned} \mathbf{P}(t_k|t_{k-1}) &= E[\mathbf{e}(t_k|t_{k-1})\mathbf{e}^T(t_k|t_{k-1})] \\ &= \mathbf{F}(t_k)\mathbf{P}(t_{k-1}|t_{k-1})\mathbf{F}^T(t_k) + \mathbf{Q}(t_k) \end{aligned} \quad (\text{C.12})$$

Der prädizierte Zustand und die Fehlerkovarianzmatrizen können anhand der Beobachtungsmatrix und per Varianzfortpflanzung in denselben Raum wie die Beobachtungen transformiert werden:

$$\hat{\mathbf{z}}(t_k|t_{k-1}) = \mathbf{H}(t_k) \cdot \hat{\mathbf{x}}(t_k|t_{k-1}) \quad (\text{C.13})$$

$$\mathbf{P}_{\mathbf{z}}(t_k|t_{k-1}) = \mathbf{H}(t_k) \cdot \mathbf{P}(t_k|t_{k-1}) \cdot \mathbf{H}^T(t_k) \quad (\text{C.14})$$

Für den Schritt der Schätzung wird zunächst die Innovation berechnet. Die entspricht der Differenz zwischen der Beobachtung und der Prädiktion und kann auch als der Prädiktionsfehler interpretiert werden:

$$\tilde{\mathbf{y}}(t_k) = \mathbf{z}(t_k) - \hat{\mathbf{z}}(t_k|t_{k-1}) = \mathbf{z}(t_k) - \mathbf{H}(t_k)\hat{\mathbf{x}}(t_k|t_{k-1}) \quad (\text{C.15})$$

Die Innovationskovarianz beträgt die Summe der Kovarianzen der beiden Vektoren im Zeitpunkt  $t_k$ :

$$\mathbf{S}(t_k) = \mathbf{P}_{\hat{\mathbf{z}}}(t_k) + \mathbf{P}_{\mathbf{z}}(t_k) = \mathbf{H}(t_k)\mathbf{P}(t_k|t_{k-1})\mathbf{H}^T(t_k) + \mathbf{R}(t_k) \quad (\text{C.16})$$

Schließlich soll der Zustandsvektor anhand der Innovation aktualisiert werden. Dabei wird die Innovation mit der Kalman-Matrix  $\mathbf{K}$  gewichtet:

$$\begin{aligned} \hat{\mathbf{x}}(t_k|t_k) &= \hat{\mathbf{x}}(t_k|t_{k-1}) + \mathbf{K}(t_k)\tilde{\mathbf{y}}(t_k) \\ &= \hat{\mathbf{x}}(t_k|t_{k-1}) + \mathbf{K}(t_k)[\mathbf{z}(t_k) - \mathbf{H}(t_k)\hat{\mathbf{x}}(t_k|t_{k-1})] \end{aligned} \quad (\text{C.17})$$

Die Kalman-Matrix wird so ausgewählt, dass die aktualisierte Schätzung optimiert wird. Zur Optimierung wird die Minimierung des mittleren quadratischen Fehlers angestrebt (dabei  $\hat{\mathbf{x}}(t_k|t_k)$  entspricht  $\hat{\mathbf{x}}(t_k)$ ).

$$\mathbf{P}(t_k|t_k) = \mathbb{E}[\mathbf{e}(t_k)\mathbf{e}^T(t_k)] = \mathbb{E}[(\tilde{\mathbf{x}}(t_k) - \hat{\mathbf{x}}(t_k))(\tilde{\mathbf{x}}(t_k) - \hat{\mathbf{x}}(t_k))^T] \quad (\text{C.18})$$

Durch Einsetzen von Gleichung (C.9) in (C.17) und dann von dieser in (C.18) erhält man:

$$\begin{aligned} \mathbf{P}(t_k|t_k) &= \mathbb{E}\{[(\tilde{\mathbf{x}}(t_k) - \hat{\mathbf{x}}(t_k|t_{k-1})) \\ &\quad - \mathbf{K}(t_k)(\mathbf{H}(t_k) \cdot \tilde{\mathbf{x}}(t_k) + \mathbf{v}_s(t_k) - \mathbf{H}(t_k)\hat{\mathbf{x}}(t_k|t_{k-1}))] \cdot \\ &\quad \cdot [(\tilde{\mathbf{x}}(t_k) - \hat{\mathbf{x}}(t_k|t_{k-1})) \\ &\quad - \mathbf{K}(t_k)(\mathbf{H}(t_k) \cdot \tilde{\mathbf{x}}(t_k) + \mathbf{v}_s(t_k) - \mathbf{H}(t_k)\hat{\mathbf{x}}(t_k|t_{k-1}))]^T\} \end{aligned} \quad (\text{C.19})$$

Unter der Annahme, dass der prädizierte Zustandsfehler und der Messfehler unkorreliert sind, ergibt sich Folgendes (mit  $\mathbf{I}$  als Identitätsmatrix):

$$\mathbf{P}(t_k|t_k) = [\mathbf{I} - \mathbf{K}(t_k)\mathbf{H}(t_k)]\mathbf{P}(t_k|t_{k-1})[\mathbf{I} - \mathbf{K}(t_k)\mathbf{H}(t_k)]^T + \mathbf{K}(t_k)\mathbf{R}(t_k)\mathbf{K}(t_k) \quad (\text{C.20})$$

Die Suche der Matrix  $\mathbf{K}$ , die diese Fehlerkovarianzmatrix minimiert, ergibt (siehe Herleitung in Brown und Hwang [23]):

$$\mathbf{K}(t_k) = \mathbf{P}(t_k|t_{k-1})\mathbf{H}^T(t_k)\mathbf{S}^{-1}(t_k) \quad (\text{C.21})$$

Unter Einsatz dieser optimalen Kalman-Matrix kann Ausdruck (C.20) für die Fehlerkovarianzmatrix so formuliert werden:

$$\mathbf{P}(t_k|t_k) = [\mathbf{I} - \mathbf{K}(t_k)\mathbf{H}(t_k)]\mathbf{P}(t_k|t_{k-1}) \quad (\text{C.22})$$

Zusammenfassend besteht damit der Kalman-Filter aus der Zustandsprädiktion mit Gleichung (C.10), der Prädiktion der Fehlerkovarianzmatrix mit Gleichung (C.12), der Berechnung der Kalman-Matrix (C.21), der Zustandsschätzung mit Gleichung (C.17) und der Schätzung der Fehlerkovarianzmatrix (C.22).

Da der Filter auf der Annahme beruht, dass die letzte Schätzung vorliegt, muss die Filterinitialisierung zusätzlich gelöst werden,  $\mathbf{x}(t_0)$  und  $\mathbf{P}(t_0)$ .

### C.3 Der Informationsfilter

Der Informationsfilter<sup>2</sup> (IF) kann als mathematische Umformulierung des Kalman-Filters gesehen werden, in der die Fisher-Information anstatt des Zustandsvektors verwendet wird. Die Fisher-Information ist ein Maß der Information über einer Zufallsgröße  $\mathbf{x}$ , die in einer Reihe von Beobachtungen  $\mathbf{Z}$  bis zum Zeitpunkt  $t_k$  beinhaltet ist. Im Folgenden wird die Herleitung des Informationsfilters angelehnt an Mutambara [109] gegeben.

Die Grundstrukturen des Informationsfilters sind die Informationsmatrix und der Informationszustandsvektor:

$$\mathbf{Y}(t_i|t_j) \triangleq \mathbf{P}^{-1}(t_i|t_j) \quad (\text{C.23})$$

$$\hat{\mathbf{y}}(t_i|t_j) \triangleq \mathbf{P}^{-1}(t_i|t_j) \cdot \hat{\mathbf{x}}(t_i|t_j) \quad (\text{C.24})$$

Der Ausdruck der Schätzung im Kalman-Filter aus Gleichung (C.17) wird nun wie folgt umformuliert:

---

<sup>2</sup>Auch als die Inverskovarianzform des Kalman-Filters.

$$\hat{\mathbf{x}}(t_k|t_k) = [\mathbf{I} - \mathbf{K}(t_k)\mathbf{H}(t_k)]\hat{\mathbf{x}}(t_k|t_{k-1}) + \mathbf{K}(t_k)\mathbf{z}(t_k) \quad (\text{C.25})$$

Daraus wird der Ausdruck  $[\mathbf{I} - \mathbf{K}(t_k)\mathbf{H}(t_k)]$  mit  $\mathbf{P}(t_k|t_{k-1})\mathbf{P}^{-1}(t_k|t_{k-1})$  von rechts multipliziert und anhand der Gleichung (C.22) so ausgedrückt:

$$\begin{aligned} \mathbf{I} - \mathbf{K}(t_k)\mathbf{H}(t_k) &= [\mathbf{I} - \mathbf{K}(t_k)\mathbf{H}(t_k)]\mathbf{P}(t_k|t_{k-1})\mathbf{P}^{-1}(t_k|t_{k-1}) \\ &= \mathbf{P}(t_k|t_k)\mathbf{P}^{-1}(t_k|t_{k-1}) \end{aligned} \quad (\text{C.26})$$

Durch Einsetzen der Innovationskovarianz aus Gleichung (C.16) in den Ausdruck der Kalman-Matrix (C.21) ergibt sich folgendes:

$$\begin{aligned} \mathbf{K}(t_k) &= \mathbf{P}(t_k|t_{k-1})\mathbf{H}^T(t_k)[\mathbf{H}(t_k)\mathbf{P}(t_k|t_{k-1})\mathbf{H}^T(t_k) + \mathbf{R}(t_k)]^{-1} \\ \mathbf{K}(t_k)[\mathbf{H}(t_k)\mathbf{P}(t_k|t_{k-1})\mathbf{H}^T(t_k) + \mathbf{R}(t_k)] &= \mathbf{P}(t_k|t_{k-1})\mathbf{H}^T(t_k) \\ \mathbf{K}(t_k)\mathbf{H}(t_k)\mathbf{P}(t_k|t_{k-1})\mathbf{H}^T(t_k) + \mathbf{K}(t_k)\mathbf{R}(t_k) &= \mathbf{P}(t_k|t_{k-1})\mathbf{H}^T(t_k) \\ \mathbf{K}(t_k)\mathbf{R}(t_k) &= [\mathbf{I} - \mathbf{K}(t_k)\mathbf{H}(t_k)]\mathbf{P}(t_k|t_{k-1})\mathbf{H}^T(t_k) \\ \mathbf{K}(t_k) &= [\mathbf{I} - \mathbf{K}(t_k)\mathbf{H}(t_k)]\mathbf{P}(t_k|t_{k-1})\mathbf{H}^T(t_k)\mathbf{R}^{-1}(t_k) \end{aligned} \quad (\text{C.27})$$

Dabei wird (C.26) eingesetzt:

$$\begin{aligned} \mathbf{K}(t_k) &= \mathbf{P}(t_k|t_k)\mathbf{P}^{-1}(t_k|t_{k-1})\mathbf{P}(t_k|t_{k-1})\mathbf{H}^T(t_k)\mathbf{R}^{-1}(t_k) \\ &= \mathbf{P}(t_k|t_k)\mathbf{H}^T(t_k)\mathbf{R}^{-1}(t_k) \end{aligned} \quad (\text{C.28})$$

Nun werden die Ausdrücke (C.26) und (C.28) in der Gleichung der Zustandsschätzung (C.25) eingesetzt und als Informationsvektor ausgedrückt:

$$\begin{aligned} \hat{\mathbf{x}}(t_k|t_k) &= \mathbf{P}(t_k|t_k)\mathbf{P}^{-1}(t_k|t_{k-1})\hat{\mathbf{x}}(t_k|t_{k-1}) + \mathbf{P}(t_k|t_k)\mathbf{H}^T(t_k)\mathbf{R}^{-1}(t_k)\mathbf{z}(t_k) \\ \mathbf{P}^{-1}(t_k|t_k)\hat{\mathbf{x}}(t_k|t_k) &= \mathbf{P}^{-1}(t_k|t_{k-1})\hat{\mathbf{x}}(t_k|t_{k-1}) + \mathbf{H}^T(t_k)\mathbf{R}^{-1}(t_k)\mathbf{z}(t_k) \\ \hat{\mathbf{y}}(t_k|t_k) &= \hat{\mathbf{y}}(t_k|t_{k-1}) + \mathbf{H}^T(t_k)\mathbf{R}^{-1}(t_k)\mathbf{z}(t_k) \end{aligned} \quad (\text{C.29})$$



Letztere entspricht der Schätzung des Informationszustandsvektors. Für die Informationsmatrix wird die Gleichung (C.20) verwendet und ebenso mit den Ausdrücken (C.26) und (C.28) ersetzt:

$$\begin{aligned} \mathbf{P}(t_k|t_k) &= [\mathbf{P}(t_k|t_k)\mathbf{P}^{-1}(t_k|t_{k-1})]\mathbf{P}(t_k|t_{k-1})[\mathbf{P}(t_k|t_k)\mathbf{P}^{-1}(t_k|t_{k-1})]^T \\ &\quad + [\mathbf{P}(t_k|t_k)\mathbf{H}^T(t_k)\mathbf{R}^{-1}(t_k)]\mathbf{R}(t_k)[\mathbf{P}(t_k|t_k)\mathbf{H}^T(t_k)\mathbf{R}^{-1}(t_k)]^T \end{aligned} \quad (\text{C.30})$$

Nach einer Links- und einer Rechts-Multiplikation mit  $\mathbf{P}^{-1}(t_k|t_k)$  und nach Vereinfachung ergibt sich die Schätzung der Informationsmatrix:

$$\begin{aligned} \mathbf{P}^{-1}(t_k|t_k) &= \mathbf{P}^{-1}(t_{k-1}|t_{k-1}) + \mathbf{H}^T(t_k)\mathbf{R}^{-1}(t_k)\mathbf{H}(t_k) \\ \mathbf{Y}(t_k|t_k) &= \mathbf{Y}(t_{k-1}|t_{k-1}) + \mathbf{H}^T(t_k)\mathbf{R}^{-1}(t_k)\mathbf{H}(t_k) \end{aligned} \quad (\text{C.31})$$

Aus Gleichungen (C.29) und (C.31) werden die Informationsbeiträge wie folgt definiert:

$$\mathbf{I}(t_i) \triangleq \mathbf{H}^T(t_i) \cdot \mathbf{R}^{-1}(t_i) \cdot \mathbf{H}(t_i) \quad (\text{C.32})$$

$$\mathbf{i}(t_i) \triangleq \mathbf{H}^T(t_i) \cdot \mathbf{R}^{-1}(t_i) \cdot \mathbf{z}(t_i) \quad (\text{C.33})$$

Und die Schätzung entspricht damit:

$$\mathbf{Y}(t_k|t_k) = \mathbf{Y}(t_k|t_{k-1}) + \mathbf{I}(t_k) \quad (\text{C.34})$$

$$\hat{\mathbf{y}}(t_k|t_k) = \hat{\mathbf{y}}(t_k|t_{k-1}) + \mathbf{i}(t_k) \quad (\text{C.35})$$

Mithilfe der folgenden Matrix

$$\mathbf{L}(t_k|t_{k-1}) = \mathbf{Y}(t_k|t_{k-1})\mathbf{F}(t_k)\mathbf{Y}(t_{k-1}|t_{k-1}) \quad (\text{C.36})$$

kann auch die Prädiktion in Form von Information bestimmt werden:

$$\hat{\mathbf{y}}(t_k|t_{k-1}) = \mathbf{L}(t_k|t_{k-1})\hat{\mathbf{y}}(t_{k-1}|t_{k-1}) \quad (\text{C.37})$$

$$\mathbf{Y}(t_k|t_{k-1}) = [\mathbf{F}(t_k)\mathbf{Y}^{-1}(t_{k-1}|t_{k-1})\mathbf{F}^T(t_k) + \mathbf{Q}(t_k)]^{-1} \quad (\text{C.38})$$

Mathematisch äquivalent ist jedoch die Prädiktion anhand des Kalman-Filters durchzuführen – Gleichungen (C.10) und (C.12) – und diese in die Informationsform zu bringen – Gleichungen (C.23) und (C.24).

Damit entsprechen die Ausdrücke des Informationsfilters der Prädiktion anhand der Gleichungen (C.10) und (C.12) sowie der Schätzung anhand der Gleichungen (C.34) und (C.35).

Wenn es sich um mehrere gleichzeitige Beobachtungen handelt, dann ergibt sich:

$$\mathbf{Y}(t_k|t_k) = \mathbf{Y}(t_k|t_{k-1}) + \sum_{\mathbf{s}} \mathbf{I}_{\mathbf{s}}(t_k) \quad (\text{C.39})$$

$$\hat{\mathbf{y}}(t_k|t_k) = \hat{\mathbf{y}}(t_k|t_{k-1}) + \sum_{\mathbf{s}} \mathbf{i}_{\mathbf{s}}(t_k) \quad (\text{C.40})$$

Die Vorteile dieses Filters werden in Abschnitt 7.5.3 diskutiert.

## C.4 Verwendbarkeit des Kalman-Filters für indirekte Beobachtungen

Im Zusammenhang mit der in Abschnitt 3.1 durchgeführten Diskussion zur Ebene der Fusion wird hier bewiesen, dass ein Kalman-Filter<sup>3</sup> auf Merkmals- und Objektebene äquivalent ist, solange die Messdaten und die Messunsicherheit entsprechend angepasst werden – siehe folgende Tabelle.

Merkmalsebene	Objektebene
$\mathbf{x}(t_k), \mathbf{P}(t_k)$	$\mathbf{x}(t_k), \mathbf{P}(t_k)$
$\mathbf{F}(t_k)$	$\mathbf{F}(t_k)$
$\mathbf{Q}(t_k)$	$\mathbf{Q}(t_k)$
$\mathbf{z}_m(t_k)$	$\mathbf{z}_o(t_k) \triangleq \mathbf{H}_t^{-1}(t_k)\mathbf{z}_m(t_k)$
$\mathbf{R}_m(t_k)$	$\mathbf{R}_o(t_k) \triangleq \mathbf{H}_t^{-1}(t_k)\mathbf{R}_m(t_k)(\mathbf{H}_t^{-1}(t_k))^T$
$\mathbf{H}_m(t_k) \triangleq \mathbf{H}_t(t_k)\mathbf{H}_o(t_k)$	$\mathbf{H}_o(t_k) \triangleq \mathbf{H}_t^{-1}(t_k)\mathbf{H}_m(t_k)$

Die Filterung auf Merkmalsebene kann wie folgt dargestellt werden:

$$\mathbf{z}_m, \mathbf{R}_m \rightarrow \boxed{\begin{matrix} \mathbf{F}, \mathbf{Q}, \mathbf{H}_m \\ \text{Filter} \end{matrix}} \rightarrow \mathbf{x}, \mathbf{P}$$

Auf Objektebene wird die Beobachtungsmatrix in einen merkmalspezifischen Teil  $\mathbf{H}_t$  (Vorverarbeitung außerhalb des Filters) und einen objektspezifischen Teil  $\mathbf{H}_o$  (Trivial) zerlegt:

$$\mathbf{z}_m, \mathbf{R}_m \rightarrow \boxed{\begin{matrix} \mathbf{H}_t \\ \text{Vorverarb.} \end{matrix}} \rightarrow \mathbf{z}_o, \mathbf{R}_o \rightarrow \boxed{\begin{matrix} \mathbf{F}, \mathbf{Q}, \mathbf{H}_o \\ \text{Filter} \end{matrix}} \rightarrow \mathbf{x}, \mathbf{P}$$

Unter der Annahme, dass der merkmalspezifische Teil  $\mathbf{H}_t$  invertierbar ist, d.h. es lassen sich Objektbeobachtungen aus den Sensordaten extrahieren, wird im Folgendem gezeigt, dass beide Filter äquivalent sind.

Der Kalman-Filter auf Objektebene (mit  $\mathbf{z}_o$ ,  $\mathbf{R}_o$  und  $\mathbf{H}_o$ ) wird als Basis genommen (Gleichungen C.6 und C.9).

$$\mathbf{x}(t_k) = \mathbf{F}(t_k)\mathbf{x}(t_{k-1}) + \mathbf{B}(t_k)\mathbf{u}(t_k) + \mathbf{w}(t_k) \quad (\text{C.41})$$

$$\mathbf{z}_o(t_k) = \mathbf{H}_o(t_k)\mathbf{x}(t_k) + \mathbf{v}_o(t_k) \quad (\text{C.42})$$

Die Zustandsprädiktion bleibt unverändert (Gleichungen C.10 und C.12):

<sup>3</sup>Äquivalent für einen Informationsfilter

$$\hat{\mathbf{x}}(t_k|t_{k-1}) = \mathbf{F}(t_k)\hat{\mathbf{x}}(t_{k-1}|t_{k-1}) \quad (\text{C.43})$$

$$\mathbf{P}(t_k|t_{k-1}) = \mathbf{F}(t_k)\mathbf{P}(t_{k-1}|t_{k-1})\mathbf{F}^T(t_k) + \mathbf{Q}(t_k) \quad (\text{C.44})$$

Die Beobachtungsgleichung (C.42) kann wie folgt anhand der letzten Tabelle (rechts) im Merkmalsraum ausgedrückt werden:

$$\mathbf{H}_{\mathbf{t}}^{-1}(t_k)\mathbf{z}_{\mathbf{m}}(t_k) = \mathbf{H}_{\mathbf{o}}(t_k)\mathbf{x}(t_k) + \mathbf{v}_{\mathbf{o}}(t_k) \quad (\text{C.45})$$

$$\mathbf{z}_{\mathbf{m}}(t_k) = \mathbf{H}_{\mathbf{t}}(t_k)\mathbf{H}_{\mathbf{o}}(t_k)\mathbf{x}(t_k) + \mathbf{H}_{\mathbf{t}}(t_k)\mathbf{v}_{\mathbf{o}}(t_k) \quad (\text{C.46})$$

$$\mathbf{z}_{\mathbf{m}}(t_k) = \mathbf{H}_{\mathbf{m}}(t_k)\mathbf{x}(t_k) + \mathbf{H}_{\mathbf{t}}(t_k)\mathbf{v}_{\mathbf{o}}(t_k) \quad (\text{C.47})$$

Die Messwertprädiktion bzw. Beobachtungsprädiktion für die Merkmalsebene (Gleichung C.13) wird analog zu:

$$\hat{\mathbf{z}}_{\mathbf{o}}(t_k|t_{k-1}) = \mathbf{H}_{\mathbf{o}}(t_k) \cdot \hat{\mathbf{x}}(t_k|t_{k-1}) \quad (\text{C.48})$$

$$\hat{\mathbf{z}}_{\mathbf{m}}(t_k|t_{k-1}) = \mathbf{H}_{\mathbf{t}}(t_k)\mathbf{H}_{\mathbf{o}}(t_k) \cdot \hat{\mathbf{x}}(t_k|t_{k-1}) \quad (\text{C.49})$$

$$\hat{\mathbf{z}}_{\mathbf{m}}(t_k|t_{k-1}) = \mathbf{H}_{\mathbf{m}}(t_k) \cdot \hat{\mathbf{x}}(t_k|t_{k-1}) \quad (\text{C.50})$$

Dieser Ausdruck entspricht erwartungsgemäß der linken Seite der Tabelle. Die prädizierte Messfehlerkovarianz (Gleichung C.14) wird anhand der Varianzfortpflanzung, unter Kenntnis dass  $\hat{\mathbf{z}}_{\mathbf{m}}(t_k) = \mathbf{H}_{\mathbf{t}}(t_k)\hat{\mathbf{z}}_{\mathbf{o}}(t_k)$  ist, zu:

$$\mathbf{P}_{\hat{\mathbf{z}}_{\mathbf{o}}}(t_k|t_{k-1}) = \mathbf{H}_{\mathbf{o}}(t_k) \cdot \mathbf{P}(t_k|t_{k-1}) \cdot \mathbf{H}_{\mathbf{o}}^T(t_k) \quad (\text{C.51})$$

$$\mathbf{P}_{\hat{\mathbf{z}}_{\mathbf{m}}}(t_k|t_{k-1}) = \mathbf{H}_{\mathbf{t}}(t_k) \cdot \mathbf{H}_{\mathbf{o}}(t_k) \cdot \mathbf{P}(t_k|t_{k-1}) \cdot \mathbf{H}_{\mathbf{o}}^T(t_k) \cdot \mathbf{H}_{\mathbf{t}}^T(t_k) \quad (\text{C.52})$$

$$\mathbf{P}_{\hat{\mathbf{z}}_{\mathbf{o}}}(t_k|t_{k-1}) = \mathbf{H}_{\mathbf{t}}^{-1}(t_k) \cdot \mathbf{P}_{\hat{\mathbf{z}}_{\mathbf{m}}}(t_k|t_{k-1}) \cdot (\mathbf{H}_{\mathbf{t}}^{-1}(t_k))^T \quad (\text{C.53})$$

Die Innovation (Gleichung C.15) im Merkmalsraum entspricht somit:

$$\tilde{\mathbf{y}}_{\mathbf{o}}(t_k) = \mathbf{z}_{\mathbf{o}}(t_k) - \mathbf{H}_{\mathbf{o}}(t_k)\hat{\mathbf{x}}(t_k|t_{k-1}) \quad (\text{C.54})$$

$$\tilde{\mathbf{y}}_{\mathbf{o}}(t_k) = \mathbf{H}_{\mathbf{t}}^{-1}(t_k)\mathbf{z}_{\mathbf{m}}(t_k) - \mathbf{H}_{\mathbf{t}}^{-1}(t_k)\hat{\mathbf{z}}_{\mathbf{m}}(t_k|t_{k-1}) \quad (\text{C.55})$$

$$\tilde{\mathbf{y}}_{\mathbf{o}}(t_k) = \mathbf{H}_{\mathbf{t}}^{-1}(t_k)\tilde{\mathbf{y}}_{\mathbf{m}}(t_k) \quad (\text{C.56})$$

Die Innovationskovarianz (aus Gleichung C.16) lautet:

$$\mathbf{S}_{\mathbf{o}}(t_k) = \mathbf{H}_{\mathbf{o}}(t_k)\mathbf{P}(t_k|t_{k-1})\mathbf{H}_{\mathbf{o}}^T(t_k) + \mathbf{R}_{\mathbf{o}}(t_k) \quad (\text{C.57})$$

$$\begin{aligned} \mathbf{S}_{\mathbf{o}}(t_k) &= \mathbf{H}_{\mathbf{t}}^{-1}(t_k) \cdot \mathbf{P}_{\hat{\mathbf{z}}_{\mathbf{m}}}(t_k|t_{k-1}) \cdot (\mathbf{H}_{\mathbf{t}}^{-1}(t_k))^T \\ &\quad + \mathbf{H}_{\mathbf{t}}^{-1}(t_k)\mathbf{R}_{\mathbf{m}}(t_k)(\mathbf{H}_{\mathbf{t}}^{-1}(t_k))^T \end{aligned} \quad (\text{C.58})$$

Da es sich um eine lineare Funktion handelt, kann folgende Umformulierung gemacht werden:

$$\mathbf{S}_{\mathbf{o}}(t_k) = \mathbf{H}_{\mathbf{t}}^{-1}(t_k) \cdot (\mathbf{P}_{\hat{\mathbf{z}}_{\mathbf{m}}}(t_k|t_{k-1}) + \mathbf{R}_{\mathbf{m}}(t_k)) \cdot (\mathbf{H}_{\mathbf{t}}^{-1}(t_k))^T \quad (\text{C.59})$$

$$\mathbf{S}_{\mathbf{o}}(t_k) = \mathbf{H}_{\mathbf{t}}^{-1}(t_k) \cdot \mathbf{S}_{\mathbf{m}}(t_k) \cdot (\mathbf{H}_{\mathbf{t}}^{-1}(t_k))^T \quad (\text{C.60})$$

Die Kalman-Matrix (aus Gleichung C.21) ergibt sich mit Einsatz von (C.60) zu:

$$\mathbf{K}_{\mathbf{o}}(t_k) = \mathbf{P}(t_k|t_{k-1})\mathbf{H}_{\mathbf{o}}^T(t_k)\mathbf{S}_{\mathbf{o}}^{-1}(t_k) \quad (\text{C.61})$$

$$\mathbf{K}_{\mathbf{o}}(t_k) = \mathbf{P}(t_k|t_{k-1})\mathbf{H}_{\mathbf{o}}^T(t_k)[\mathbf{H}_{\mathbf{t}}^{-1}(t_k)\mathbf{S}_{\mathbf{m}}(t_k)(\mathbf{H}_{\mathbf{t}}^{-1}(t_k))^T]^{-1} \quad (\text{C.62})$$

$$\mathbf{K}_{\mathbf{o}}(t_k) = \mathbf{P}(t_k|t_{k-1})\mathbf{H}_{\mathbf{o}}^T(t_k)\mathbf{H}_{\mathbf{t}}^T(t_k)\mathbf{S}_{\mathbf{m}}^{-1}(t_k)\mathbf{H}_{\mathbf{t}}(t_k) \quad (\text{C.63})$$

$$\mathbf{K}_{\mathbf{o}}(t_k) = \mathbf{P}(t_k|t_{k-1})[\mathbf{H}_{\mathbf{t}}(t_k)\mathbf{H}_{\mathbf{o}}(t_k)]^T\mathbf{S}_{\mathbf{m}}^{-1}(t_k)\mathbf{H}_{\mathbf{t}}(t_k) \quad (\text{C.64})$$

$$\mathbf{K}_{\mathbf{o}}(t_k) = \mathbf{K}_{\mathbf{m}}(t_k)\mathbf{H}_{\mathbf{t}}(t_k) \quad (\text{C.65})$$

Die Schätzung (C.17) wird dann anhand der Gleichungen (C.56) und (C.65) wie folgt umgestellt:

$$\hat{\mathbf{x}}(t_k|t_k) = \hat{\mathbf{x}}(t_k|t_{k-1}) + \mathbf{K}_o(t_k)\tilde{\mathbf{y}}_o(t_k) \quad (\text{C.66})$$

$$\hat{\mathbf{x}}(t_k|t_k) = \hat{\mathbf{x}}(t_k|t_{k-1}) + \mathbf{K}_m(t_k)\mathbf{H}_t(t_k)\mathbf{H}_t^{-1}(t_k)\tilde{\mathbf{y}}_m(t_k) \quad (\text{C.67})$$

$$\hat{\mathbf{x}}(t_k|t_k) = \hat{\mathbf{x}}(t_k|t_{k-1}) + \mathbf{K}_m(t_k)\tilde{\mathbf{y}}_m(t_k) \quad (\text{C.68})$$

Letztere entspricht dem Ausdruck auf Merkmalsebene. Analog für die Schätzung der Kovarianzmatrix (C.22) erhält man folgende Ausdrücke auf Merkmalsebene:

$$\mathbf{P}(t_k|t_k) = [\mathbf{I} - \mathbf{K}_o(t_k)\mathbf{H}_o(t_k)]\mathbf{P}(t_k|t_{k-1}) \quad (\text{C.69})$$

$$\mathbf{P}(t_k|t_k) = [\mathbf{I} - \mathbf{K}_m(t_k)\mathbf{H}_t(t_k)\mathbf{H}_o(t_k)]\mathbf{P}(t_k|t_{k-1}) \quad (\text{C.70})$$

$$\mathbf{P}(t_k|t_k) = [\mathbf{I} - \mathbf{K}_m(t_k)\mathbf{H}_m(t_k)]\mathbf{P}(t_k|t_{k-1}) \quad (\text{C.71})$$

Damit gilt die Prüfung, dass die verwendeten Ausdrücke auf Objektebene denjenigen der Merkmalsebene entsprechen. Dafür ist es notwendig, den Messvektor und die Messrauschenmatrix entsprechend der gegebenen Tabelle anzupassen.

## D Beispiele der partiellen Beobachtbarkeit

Die Ausdrücke in diesem Anhang sind in Abschnitt 7.6 eingeführt worden.

In diesem Anhang werden einige Beispiele vorgestellt, die die Bestimmung der partiellen Beobachtbarkeit der Komponenten des Zustandsvektors erläutern. Diese umfassen Fälle mit direkter Beobachtbarkeit, mit indirekter Beobachtbarkeit und mit Beobachtbarkeit aus mehreren Messkomponenten. Für jedes Beispiel wird das System ( $\mathbf{x}(t_k)$  und  $\mathbf{F}(t_k)$ ,  $\mathbf{z}(t_k)$  und  $\mathbf{H}(t_k)$ ) beschrieben (a). Daraus wird die Matrix  $\mathbf{M}$  aus Gleichung (7.25) auf Seite 79 hergeleitet und anhand ihres Rangs eine Aussage über die Beobachtbarkeit des Systems getroffen (b). Des Weiteren wird die partielle Beobachtbarkeit der Beobachtungen eingeschränkt und daraus die Systembeobachtbarkeit (für jede Komponente) hergeleitet (c). Schließlich wird die Zustandsprädiktion und die Messwertprädiktion anhand der eingeschränkten Beobachtbarkeit hergeleitet (d).

### Beispiel 1: direkte Beobachtbarkeit

a) Ein System sei durch seine Position gegeben und in Form von folgendem Zustandsvektor und Übergangsmatrix definiert:

$$\mathbf{x}(t_k) = \begin{bmatrix} x \\ y \end{bmatrix}; \mathbf{F}(t_k) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (\text{D.1})$$

Wenn ein Sensor beide Komponenten erfassen kann, lauten der Messvektor und die Beobachtungsmatrix:

$$\mathbf{z}(t_k) = \begin{bmatrix} x \\ y \end{bmatrix}; \mathbf{H}(t_k) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (\text{D.2})$$

b) Für dieses System ergibt sich die Matrix  $\mathbf{M}$  aus Gleichung (7.25) auf Seite 79 zu:

$$\mathbf{M} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (\text{D.3})$$

Ihr Rang beträgt 2, wie die Dimension des Zustandsvektors  $n_x = 2$ . Daher ist das System vollständig beobachtbar.

c) Die partielle Beobachtbarkeit hängt jedoch von der Beobachtbarkeit der eingehenden Beobachtungen ab. Ist beispielsweise bei einer Beobachtung nur die Komponente  $x$  beobachtbar

$$\mathbf{V}_z(t_k) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad (\text{D.4})$$

so entspricht die angepasste Matrix  $\mathbf{M}_V$  aus Gleichung (7.38) auf Seite 82:

$$\mathbf{M}_V = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \quad (\text{D.5})$$

Ihr Rang beträgt  $\text{Rang}(\mathbf{M}_V) = 1$ . So lässt sich die partielle Beobachtbarkeit für jede Komponente wie folgt bestimmen:

$$v_{x_x} = \text{Rang}(\mathbf{M}_V) - \text{Rang}(\mathbf{M}_V^x) = 1 - 0 = 1; \text{ mit } \mathbf{M}_V^x = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (\text{D.6})$$

$$v_{x_y} = \text{Rang}(\mathbf{M}_V) - \text{Rang}(\mathbf{M}_V^y) = 1 - 1 = 0; \text{ mit } \mathbf{M}_V^y = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \quad (\text{D.7})$$

Und dies wird in folgender Matrix zusammengefasst:

$$\mathbf{V}_x(t_k) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad (\text{D.8})$$

d) Wird also der Zustandsvektor auf den nächsten Zeitpunkt prädiziert, so erhält man Folgendes:

$$\hat{\mathbf{x}}(t_{k+1}|t_k) = \begin{bmatrix} \hat{x} \\ 0 \end{bmatrix}; \mathbf{P}(t_{k+1}|t_k) = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & 0 \end{bmatrix} \quad (\text{D.9})$$

Für die Beobachtbarkeit des prädizierten Messvektors gilt:



$$\mathbf{D} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (\text{D.10})$$

$$v_{z_x} = 1, \sum_j d_{s,1j} = 1 + 0 = 1 \cap \sum_j d_{s,1j} v_{x_j} = 1 \cdot 1 + 0 \cdot 0 = 1 \quad (\text{D.11})$$

$$v_{z_y} = 0, \sum_j d_{s,2j} = 0 + 1 = 1 \cap \sum_j d_{s,2j} v_{x_j} = 0 \cdot 1 + 1 \cdot 0 = 0 \quad (\text{D.12})$$

Damit lauten der prädizierte Messvektor und dessen Kovarianzmatrix wie folgt:

$$\hat{\mathbf{z}}(t_{k+1}|t_k) = \begin{bmatrix} \hat{x} \\ 0 \end{bmatrix}; \mathbf{P}_{\hat{\mathbf{z}}}(t_{k+1}|t_k) = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & 0 \end{bmatrix} \quad (\text{D.13})$$

## Beispiel 2: indirekte Beobachtbarkeit

a) Sei ein System durch seine Position und Geschwindigkeit in Form von folgendem Zustandsvektor und Übergangsmatrix definiert:

$$\mathbf{x}(t_i) = \begin{bmatrix} x \\ \dot{x} \end{bmatrix}; \mathbf{F}(t_i) = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \quad (\text{D.14})$$

Wenn ein Sensor die  $x$ -Komponente erfassen kann, lauten der Messvektor und die Beobachtungsmatrix:

$$\mathbf{z}(t_i) = \begin{bmatrix} x \end{bmatrix}; \mathbf{H}(t_i) = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad (\text{D.15})$$

b) Für dieses System ergibt sich die Matrix  $\mathbf{M}$  aus Gleichung (7.25) auf Seite 79 zu:

$$\mathbf{M} = \begin{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & \Delta t \end{bmatrix} \quad (\text{D.16})$$

Ihr Rang beträgt 2, wie die Dimension des Zustandsvektors  $n_x = 2$ . Daher ist das System vollständig beobachtbar.

c) Die partielle Beobachtbarkeit hängt jedoch von der Beobachtbarkeit der eingehenden Beobachtungen ab. Ist beispielsweise die Komponente  $x$  beobachtbar, so ergibt sich

$$\mathbf{V}_z(t_i) = \begin{bmatrix} 1 \end{bmatrix} \quad (\text{D.17})$$

So entspricht die angepasste Matrix  $\mathbf{M}_V$  aus Gleichung (7.38) auf Seite 82:

$$\mathbf{M}_V = \begin{bmatrix} 1 & 0 \\ 1 & \Delta t \end{bmatrix} \quad (\text{D.18})$$

Ihr Rang beträgt  $\text{Rang}(\mathbf{M}_V) = 2$ . So lässt sich die partielle Beobachtbarkeit für jede Komponente wie folgt bestimmen:

$$v_{x_x} = \text{Rang}(\mathbf{M}_V) - \text{Rang}(\mathbf{M}_V^x) = 2 - 1 = 1; \text{ mit } \mathbf{M}_V^x = \begin{bmatrix} 0 & 0 \\ 0 & \Delta t \end{bmatrix} \quad (\text{D.19})$$

$$v_{x_{v_x}} = \text{Rang}(\mathbf{M}_V) - \text{Rang}(\mathbf{M}_V^{v_x}) = 2 - 1 = 1; \text{ mit } \mathbf{M}_V^{v_x} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad (\text{D.20})$$

Und dies wird in folgender Matrix zusammengefasst:

$$\mathbf{V}_x(t_i) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (\text{D.21})$$

d) Wird also der Zustandsvektor auf den nächsten Zeitpunkt prädiziert, so erhält man:

$$\hat{\mathbf{x}}(t_{k+1}|t_k) = \begin{bmatrix} \hat{x} \\ \dot{\hat{x}} \end{bmatrix}; \mathbf{P}(t_{k+1}|t_k) = \begin{bmatrix} \sigma_x^2 & \sigma_{x,\dot{x}}^2 \\ \sigma_{x,\dot{x}}^2 & \sigma_{\dot{x}}^2 \end{bmatrix} \quad (\text{D.22})$$

Für die Beobachtbarkeit des prädizierten Messvektors gilt:

$$\mathbf{D} = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad (\text{D.23})$$

$$v_{z_x} = 1, \sum_j d_{s,1j} = 1 + 0 = 1 \cap \sum_j d_{s,1j} v_{x_j} = 1 \cdot 1 + 0 \cdot 1 = 1 \quad (\text{D.24})$$

Damit lauten der prädizierte Messvektor und dessen Kovarianzmatrix wie folgt:

$$\hat{\mathbf{z}}(t_{k+1}|t_k) = [\hat{x}] ; \mathbf{P}_{\hat{\mathbf{z}}}(t_{k+1}|t_k) = [\sigma_x^2] \quad (\text{D.25})$$

### Beispiel 3: Beobachtbarkeit aus mehreren Messkomponenten

a) Sei ein System durch den Mittelpunkt und eine Breite definiert:

$$\mathbf{x}(t_k) = \begin{bmatrix} y \\ w \end{bmatrix} ; \mathbf{F}(t_k) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (\text{D.26})$$

Wenn ein Sensor die linke und rechte Ecke erfassen kann, lauten der Messvektor und die Beobachtungsmatrix:

$$\mathbf{z}(t_k) = \begin{bmatrix} y_l \\ y_r \end{bmatrix} ; \mathbf{H}(t_k) = \begin{bmatrix} 1 & +\frac{1}{2} \\ 1 & -\frac{1}{2} \end{bmatrix} \quad (\text{D.27})$$

b) Für dieses System ergibt sich die Matrix  $\mathbf{M}$  aus Gleichung (7.25) auf Seite 79 zu:

$$\mathbf{M} = \begin{bmatrix} \begin{bmatrix} 1 & +\frac{1}{2} \\ 1 & -\frac{1}{2} \end{bmatrix} \\ \begin{bmatrix} 1 & +\frac{1}{2} \\ 1 & -\frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 1 & +\frac{1}{2} \\ 1 & -\frac{1}{2} \\ 1 & +\frac{1}{2} \\ 1 & -\frac{1}{2} \end{bmatrix} \quad (\text{D.28})$$

Ihr Rang beträgt 2, wie die Dimension des Zustandsvektors  $n_x = 2$ . Daher ist das System vollständig beobachtbar.

c) Die partielle Beobachtbarkeit hängt jedoch von der Beobachtbarkeit der eingehenden Beobachtungen ab. Als erstes Beispiel ist nur eine Ecke beobachtbar, so dass sich ergibt:

$$\mathbf{V}_{\mathbf{z}}(t_k) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad (\text{D.29})$$

So entspricht die angepasste Matrix  $\mathbf{M}_{\mathbf{V}}$  aus Gleichung (7.38) auf Seite 82:

$$\mathbf{M}_{\mathbf{V}} = \left[ \begin{array}{c} \left[ \begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array} \right] \cdot \left[ \begin{array}{cc} 1 & +\frac{1}{2} \\ 1 & -\frac{1}{2} \end{array} \right] \\ \left[ \begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array} \right] \cdot \left[ \begin{array}{cc} 1 & +\frac{1}{2} \\ 1 & -\frac{1}{2} \end{array} \right] \cdot \left[ \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right] \end{array} \right] = \left[ \begin{array}{cc} 1 & +\frac{1}{2} \\ 0 & 0 \\ 1 & +\frac{1}{2} \\ 0 & 0 \end{array} \right] \quad (\text{D.30})$$

Ihr Rang beträgt  $\text{Rang}(\mathbf{M}_{\mathbf{V}}) = 1$ . So lässt sich die partielle Beobachtbarkeit für jede Komponente wie folgt bestimmen:

$$v_{x_y} = \text{Rang}(\mathbf{M}_{\mathbf{V}}) - \text{Rang}(\mathbf{M}_{\mathbf{V}}^y) = 1 - 1 = 0; \text{ mit } \mathbf{M}_{\mathbf{V}}^y = \left[ \begin{array}{cc} 0 & +\frac{1}{2} \\ 0 & 0 \\ 0 & +\frac{1}{2} \\ 0 & 0 \end{array} \right] \quad (\text{D.31})$$

$$v_{x_w} = \text{Rang}(\mathbf{M}_{\mathbf{V}}) - \text{Rang}(\mathbf{M}_{\mathbf{V}}^w) = 1 - 1 = 0; \text{ mit } \mathbf{M}_{\mathbf{V}}^w = \left[ \begin{array}{cc} 1 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{array} \right] \quad (\text{D.32})$$

Und dies wird in folgender Matrix für diese Beobachtung zusammengefasst:

$$\mathbf{V}_{\mathbf{x}}(t_k) = \left[ \begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right] \quad (\text{D.33})$$

Da beide Komponenten des Zustandsvektors vom ganzen Messvektor abhängen, ergibt die Nichtbeobachtbarkeit einer der Messkomponenten, dass keine Zustandskomponente beobachtet werden kann.

d.1) Wird der Zustandsvektor auf den nächsten Zeitpunkt prädiziert, so erhält man:

$$\hat{\mathbf{x}}(t_{k+1}|t_k) = \left[ \begin{array}{c} 0 \\ 0 \end{array} \right]; \mathbf{P}(t_{k+1}|t_k) = \left[ \begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right] \quad (\text{D.34})$$

Für die Beobachtbarkeit des prädizierten Messvektors gilt:

$$\mathbf{D} = \left[ \begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array} \right] \quad (\text{D.35})$$

$$v_{zy_l} = 0, \sum_j d_{s,1j} = 1 + 1 = 2 \cap \sum_j d_{s,1j} v_{x_j} = 1 \cdot 0 + 1 \cdot 0 = 0 \quad (\text{D.36})$$

$$v_{zy_r} = 0, \sum_j d_{s,1j} = 1 + 1 = 2 \cap \sum_j d_{s,1j} v_{x_j} = 1 \cdot 0 + 1 \cdot 0 = 0 \quad (\text{D.37})$$

Damit lauten der prädizierte Messvektor und dessen Kovarianzmatrix wie erwartet Folgendes:

$$\hat{\mathbf{z}}(t_{k+1}|t_k) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}; \mathbf{P}_{\hat{\mathbf{z}}}(t_{k+1}|t_k) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (\text{D.38})$$

d.2) Ist als zweites Beispiel die Position  $y$  des Objektes durch einen anderen Sensor bekannt, so kann trotzdem der Messvektor für diesen Sensor nicht prädiziert werden, da beide Komponenten vom selben Sensor stammen müssen:

$$\mathbf{V}_{\mathbf{x}}(t_k) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad (\text{D.39})$$

$$\hat{\mathbf{x}}(t_{k+1}|t_k) = \begin{bmatrix} \hat{y} \\ 0 \end{bmatrix}; \mathbf{P}(t_{k+1}|t_k) = \begin{bmatrix} \sigma_y^2 & 0 \\ 0 & 0 \end{bmatrix} \quad (\text{D.40})$$

$$\mathbf{D} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad (\text{D.41})$$

$$v_{zy_l} = 0, \sum_j d_{s,1j} = 1 + 1 = 2 \cap \sum_j d_{s,1j} v_{x_j} = 1 \cdot 1 + 1 \cdot 0 = 1 \quad (\text{D.42})$$

$$v_{zy_r} = 0, \sum_j d_{s,1j} = 1 + 1 = 2 \cap \sum_j d_{s,1j} v_{x_j} = 1 \cdot 1 + 1 \cdot 0 = 1 \quad (\text{D.43})$$

$$\hat{\mathbf{z}}(t_{k+1}|t_k) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}; \mathbf{P}_{\hat{\mathbf{z}}}(t_{k+1}|t_k) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (\text{D.44})$$

d.3) Schließlich wird hier der Fall betrachtet, in dem beide Zustandskomponenten beobachtbar

sind, womit der prädizierte Messvektor ebenso beobachtbar ist:

$$\mathbf{V}_x(t_k) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (\text{D.45})$$

$$\hat{\mathbf{x}}(t_{k+1}|t_k) = \begin{bmatrix} \hat{y} \\ \hat{w} \end{bmatrix}; \mathbf{P}(t_{k+1}|t_k) = \begin{bmatrix} \sigma_y^2 & \sigma_{y,w}^2 \\ \sigma_{y,w}^2 & \sigma_w^2 \end{bmatrix} \quad (\text{D.46})$$

$$\mathbf{D} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad (\text{D.47})$$

$$v_{z_{y_l}} = 1, \sum_j d_{s,1j} = 1 + 1 = 2 \cap \sum_j d_{s,1j} v_{x_j} = 1 \cdot 1 + 1 \cdot 1 = 2 \quad (\text{D.48})$$

$$v_{z_{y_r}} = 1, \sum_j d_{s,1j} = 1 + 1 = 2 \cap \sum_j d_{s,1j} v_{x_j} = 1 \cdot 1 + 1 \cdot 1 = 2 \quad (\text{D.49})$$

$$\hat{\mathbf{z}}(t_{k+1}|t_k) = \begin{bmatrix} \hat{y}_l \\ \hat{y}_r \end{bmatrix}; \mathbf{P}_{\hat{\mathbf{z}}}(t_{k+1}|t_k) = \begin{bmatrix} \sigma_{y_l}^2 & \sigma_{y_l, y_r}^2 \\ \sigma_{y_l, y_r}^2 & \sigma_{y_r}^2 \end{bmatrix} \quad (\text{D.50})$$

Diese Beispiele zeigen, wie die in Abschnitt 7.6 vorgestellten Methoden für die Schätzung der partiellen Beobachtbarkeit geeignet sind.

## Verwendete Abkürzungen und Formelzeichen

<i>f.n.</i>	<i>False Negative</i> (verpasstes Objekt)
<i>f.p.</i>	<i>False Positive</i> (Phantomobjekt)
ABS	Antiblockiersystem
ACC	<i>Adaptive Cruise Control</i>
DLT	<i>Direct Linear Transformation</i>
EIF	Erweiterter Informationsfilter
EKF	Erweiterter Kalman-Filter
ESP	Elektronisches Stabilitätsprogramm
IF	Informationsfilter
IMM	<i>Interacting Multi-Model filter</i>
Institut TS	Institut für Verkehrssystemtechnik <i>Institute of Transportation Systems</i>
JIPDA	<i>Joint Integrated Probabilistic Data Association</i>
KF	Kalman-Filter
LDW	<i>Lane Departure Warning</i>
LRR	<i>Long Range Radar</i>
NBA	Notbremsassistent
NIR	Near Infrared Sensor
NIS	<i>Normalized Innovation Square</i>
PMD	<i>Predicted Minimum Distance</i>
RA	Realdatenausschnitt zum Testen
SMC	Sequenzielle Monte-Carlo-Methode
SRR	<i>Short Range Radar</i>
THW	<i>Time Headway</i>
TPMD	<i>Time to Predicted Minimum Distance</i>
TS	Simuliertes Testszenario zum Testen
TTC	<i>Time To Collision</i>
UKF	<i>Unscented</i> Kalman-Filter
UTM	<i>Universal Transverse Mercator</i> Koordinatensystem
$\alpha, \alpha'$	Schwellwerte für das Gating bzw. für die Gefahrenerkennung

$\mathbf{p}_{\mathbf{c}i}$ .....	Innerer Konturpunkt eines Laserscannerobjekts
$B$ .....	Referenz eines Objekts in der Mitte der hinteren Seite
$BL$ .....	Referenz eines Objekts in der hinteren linken Ecke
$C$ .....	Referenz eines Objekts im Mittelpunkt
$F$ .....	Referenz eines Objekts in der Mitte der vorderen Seite
$FR$ .....	Referenz eines Objekts in der vorderen rechten Ecke
$L$ .....	Referenz eines Objekts in der Mitte der linken Seite
$LF$ .....	Referenz eines Objekts in der linken vorderen Ecke
$N_x$ .....	Anzahl der Objekthypothesen
$N_z$ .....	Anzahl der eingehenden Messungen eines Zyklus in der Assoziation
$R$ .....	Referenz eines Objekts in der Mitte der rechten Seite
$RB$ .....	Referenz eines Objekts in der rechten hinteren Ecke
$\Lambda(t_k)$ .....	Zusammenhang zwischen Störvektor und Prozessrauschenvektor
$\mathbf{A}$ .....	Design-Matrix eines Systems
$\mathbf{B}(t_k)$ .....	Steuereingangsmatrix eines Systems
$\mathbf{F}(t_k)$ .....	Systemmatrix
$\mathbf{H}_s(t_k)$ .....	Beobachtungsmatrix des Sensors $s$ zum Zeitpunkt $t_k$
$\mathbf{H}(t_k)$ .....	Beobachtungsmatrix zum Zeitpunkt $t_k$
$\mathbf{I}_s(t_k)$ .....	Beitrag der Beobachtung zur Informationsmatrix
$\mathbf{J}$ .....	Jacobi-Matrix eines Systems
$\mathbf{P}(t_k)$ .....	Fehlerkovarianzmatrix eines geschätzten Zustandsvektors
$\mathbf{Q}'(t_k)$ .....	Kovarianzmatrix des indirekten Prozessrauschens
$\mathbf{Q}(t_k)$ .....	Kovarianzmatrix des weißen Prozessrauschens eines Systems
$\mathbf{R}(t_k)$ .....	Kovarianzmatrix des weißen Messrauschens eines Sensors
$\mathbf{T}_c^i$ .....	Innere Transformationsmatrix (von Kamera- in Bildkoordinaten)
$\mathbf{T}_i^r$ .....	Transformationsmatrix zwischen Bild- und Straßenkoordinaten
$\mathbf{T}_{\text{rot}}$ .....	Rotationsmatrix
$\mathbf{T}_r^i$ .....	Transformationsmatrix zwischen Straßen- und Bildkoordinaten
$\mathbf{T}_{\text{trans}}$ .....	Translationsmatrix
$\mathbf{T}_v^c$ .....	Alignment Transformationsmatrix (von Fahrzeug- in Kamerakoordinaten)
$\mathbf{T}_v^i$ .....	Transformationsmatrix zwischen Fahrzeug- und Bildkoordinaten



$\mathbf{T}_w^v$ .....	Äußere Transformationsmatrix (von Welt- in Fahrzeugkoordinaten)
$\mathbf{V}_x(t_k)$ .....	Beobachtbarkeitsmatrix des Zustandsvektors
$\mathbf{V}_z(t_k)$ .....	Beobachtbarkeitsmatrix des Messvektors
$\mathbf{Y}(t_k)$ .....	Informationsmatrix
$\sigma_x$ .....	Kovarianzmatrix eines Vektors $\mathbf{x}$
$\mathcal{N}(\mu, \sigma)$ .....	Normalverteilung mit $\mu$ und $\sigma$
<i>conf</i> .....	<i>Confidence</i> bzw. Zuverlässigkeit einer Objekthypothese
$c_l, c_r$ .....	Linke und rechte Begrenzungen des Fahrkorridors
$c_{et}, \overline{c_{et}}$ .....	Korrektur der Objektposition zwischen Textur- und Kantenbild für ein Pixel und für ein Objekt
$c_{l,near}, c_{l,far}$ .....	Bestimmter möglicher Bereich der linken Fahrkorridorgrenze
$c_{r,far}, c_{r,near}$ .....	Bestimmter möglicher Bereich der rechten Fahrkorridorgrenze
$d$ .....	Metrik
$d\omega, d\phi, d\kappa$ .....	Abweichungen in der Orientierung der Kamera bzw. des Fahrzeugs zur Straße
$d_1, d_2$ .....	Dimensionen eines Laserscannerobjekts
$d_E$ .....	Metrik nach euklidischem Abstand
$d_l, d_r$ .....	Aus Laserscannerdaten bestimmte Fahrkorridorbegrenzungen
$d_M$ .....	Metrik nach Mahalanobis-Distanz
$E(x)$ .....	Erwartungswert einer Zufallsgröße
$g_{max}, g_{min}$ .....	Maximaler bzw. minimaler Grauwert in der Referenzbildregion zur Berechnung der Texturparameter
$g_X, g_Y$ .....	Gradienteninformation eines Pixels
$k_1, k_2, k_3$ .....	Einflussgrößen für den Zustandsautomaten einer Objekthypothese
$l_l, l_r$ .....	Linke und rechte Grenzen des Fahrstreifens
$n_x$ .....	Dimension des Zustandsvektors eines Systems
$n_z$ .....	Dimension eines Messvektors
$n_l$ .....	Anzahl der Fahrstreifen der Straße
$p_x$ .....	Parameter zur Datenverarbeitung
$r(t_k)$ .....	Logische Referenz eines Objekts
$st$ .....	Zustand eines Objekts
$t_k$ .....	Zeitpunkt
$t_x, t_y, t_z$ .....	Alignment-Parameter, Position des Sensors in Fahrzeugkoordinaten

$w$	Fahrstreifenbreite
$idch$	Wechsel der Objekt-Id in der Verfolgung
$\omega, \phi, \kappa$	Alignment-Parameter, Rotation um die $X$ -, $Y$ - und $Z$ -Achsen
$\bar{\sigma}_{x_i}$	Mittlere geschätzte Standardabweichung in einer Sequenz
$\bar{\varepsilon}_{x_i}$	Mittlerer absoluter Fehler in einer Sequenz
$\rho$	Entfernung zu einem Messpunkt aus dem Laserscanner
$\theta$	Winkel zum Messpunkt in einem Laserscanner
$\theta_c$	Nullausrichtung des Laserscanners
$\theta_i, \rho_i$	Parametrische Darstellung einer Gerade im Bild
$\theta_r, \rho_r$	Parametrische Darstellung einer Gerade in Straßenkoordinaten
$v_i$	Innerer Winkel eines Konturpunktes bei Laserscanner-Daten
$\hat{\mathbf{x}}$	Schätzung des Zustandsvektors
$\hat{\mathbf{z}}$	Schätzung des Beobachtungsvektors
$\mathbf{a}$	Parametervektor in der Gauß-Newton'schen Suche
$\mathbf{i}_s(t_k)$	Beitrag der Beobachtung zum Informationsvektor
$\mathbf{l}_i = [A_i \ B_i \ C_i]^T$	Gerade in homogenen Bildkoordinaten
$\mathbf{p}_l$	Punkt in Laserscanner-Koordinaten
$\mathbf{p}_p = [\theta \ \rho]$	Punkt in Polarkoordinaten
$\mathbf{p}_c$	Punkt in Kamerakoordinaten
$\mathbf{p}_{E1}, \mathbf{p}_{E2}$	Endkonturpunkte eines Laserscannerobjekts
$\mathbf{p}_i = [x_i \ y_i \ 1]^T$	Punkt in Bildkoordinaten
$\mathbf{p}_r = [x_r \ y_r \ 1]^T$	Punkt in Straßenkoordinaten
$\mathbf{p}_v = [x_v \ y_v \ z_v \ 1]^T$	Punkt in Fahrzeugkoordinaten
$\mathbf{p}_w$	Weltpunkt
$\mathbf{r}$	Residuenvektor in der Gauß-Newton'schen Suche
$\mathbf{u}(t_k)$	Steuereingang eines Systems
$\mathbf{v}(t_k)$	Messrauschenvektor eines Sensors
$\mathbf{w}'(t_k)$	Störvektor als indirektes Prozessrauschen
$\mathbf{w}(t_k)$	Prozessrauschenvektor eines Systems
$\mathbf{x}(t_k)$	Zustandsvektor eines Systems zum Zeitpunkt $t_k$
$\mathbf{y}(t_k)$	Informationszustandsvektor
$\mathbf{z}_{cam,i}$	Beobachtung aus der Bildverarbeitung in Bildkoordinaten

$\mathbf{z}_{ls}$ .....	Beobachtung aus dem Laserscanner
$\mathbf{z}_s(t_k)$ .....	Beobachtung aus Sensor $s$ zum Zeitpunkt $t_k$
$\mathbf{z}(t_k)$ .....	Messvektor zum Zeitpunkt $t_k$



# Literaturverzeichnis

- [1] *INVENT - Intelligenter Verkehr und nutzergerechte Technik*: <http://www.invent-online.de/>, geprüft am 7. Juni 2010.
- [2] *PReVENT :: INTERSAFE*: <http://www.prevent-ip.org>, geprüft am 7. Juni 2010, 2006.
- [3] *eSafetySupport.org eSafety Systems*: <http://www.esafetysupport.org/en/welcome.htm>, geprüft am 7. Juni 2010, 2007.
- [4] *MAN Nutzfahrzeuge Gruppe*: <http://www.man-mn.com/de/innovationundkompetenz>, geprüft am 7. Juni 2010, 2007.
- [5] Achler, O. und M. M. Trivedi: *Camera Based Vehicle Detection, Tracking and Wheel baseline Estimation Approach*. In: *IEEE Intelligent Transportation Systems*, S. 743–748, Piscataway, NJ, 2004. IEEE Operations Center.
- [6] Achler, O. und M. M. Trivedi: *Vehicle Wheel Detector using 2D Filter Banks*. In: *IEEE Intelligent Vehicles Symposium*, S. 25–30, Parma, Italy, 2004.
- [7] Ando, S.: *Consistent gradient operators*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(3):252–265, 2000.
- [8] Armingol, J. M., de la Escalera, Arturo, C. Hilario, J. M. Collado, J. P. Carrasco, M. J. Flores, J. M. Pastor und F. J. Rodríguez: *IVVI: Intelligent vehicle based on visual information: Robotics and Autonomous Systems in the 50th Anniversary of Artificial Intelligence, Campus Multidisciplinary in Perception and Intelligence*. *Robotics and Autonomous Systems*, 55(12):904–916, 2007.
- [9] Aycard, O.: *Grid based Fusion*. In: Polychronopoulos, A. (Hrsg.): *PReVENT Fusion Forum e-Journal*, S. 4–5. 2006.
- [10] Bar-Shalom, Y. und X.-R. Li: *Estimation and tracking: Principles, techniques, and software*. Artech House radar library. 1993.
- [11] Bar-Shalom, Y. und X.-R. Li: *Multitarget-multisensor tracking: Principles and techniques*. Yaakov Bar-Shalom, Storrs, Conn., 3rd printing Aufl., 1995.
- [12] Beauvais, M. und S. Lakshmanan: *CLARK: a heterogeneous sensor fusion method for finding lanes and obstacles*. *Image and Vision Computing*, 18(5):397–413, 2000.
- [13] Becker, J.-C.: *Fusion der Daten der objekterkennenden Sensoren eines autonomen Straßenfahrzeugs: Dissertation*, Bd. 948 d. Reihe Fortschritt-Berichte VDI Reihe 8 Meß-, Steuerungs- und Regelungstechnik. VDI-Verlag, Düsseldorf, 2002.
- [14] Benschrair, A., M. Bertozzi, A. Broggi, P. Miche, S. Mousset und G. Toulminet: *A co-operative approach to vision-based vehicle detection*. In: *IEEE Intelligent Transportation Systems*, S. 209–214, 2001.
- [15] Berger, K., C. Lipski, C. Linz, T. Stich und M. Magnor: *The Area Processing Unit of*

- Caroline - *Finding the Way through DARPA's Urban Challenge*. In: Sommer, G. und R. Klette (Hrsg.): *Robot vision*, Bd. 4931 d. Reihe *Lecture notes in computer science*, S. 260–274, Berlin, 2008. Springer-Verlag.
- [16] Bertozzi, M. und A. Broggi: *GOLD: A Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection*. IEEE Transactions On Image Processing, 7(1):62–81, 1998.
- [17] Bombini, L., P. Cerri, P. Medici und G. Alessandretti: *Radar-vision fusion for vehicle detection*. In: *3rd International Workshop on Intelligent Transportation*, Hamburg, 2006. WIT-Workshop Organization.
- [18] Bradski, G. R. und A. Kaehler: *Learning OpenCV: computer vision with the OpenCV library*. Software that sees. O'Reilly, Beijing, 1 Aufl., 2008.
- [19] Brandt, T., T. Sattel und J. Wallaschek: *Ein Ansatz zur automatischen Kollisionsvermeidung: On Automatic Collision Avoidance*. In: *Integrierte Sicherheit und Fahrerassistenzsysteme*, S. 125–143. VDI-Verlag, 2004.
- [20] Broadhurst, A., S. Baker und T. Kanade: *Monte Carlo road safety reasoning*. In: *IEEE Intelligent Vehicles Symposium*, S. 319–324, 2005.
- [21] Brooks, R. R. und S. S. Iyengar: *Multi-sensor fusion: Fundamentals and Applications with Software*. Prentice Hall PTR, Upper Saddle River, NJ, 1998.
- [22] Brown, D. C.: *Close-Range Camera Calibration*. Photogrammetric engineering, 37(8):855–866, 1971.
- [23] Brown, R. G. und P. Y. C. Hwang: *Introduction to random signals and applied Kalman filtering: With MATLAB exercises and solutions*. Wiley, New York, 3 Aufl., 1997.
- [24] Bundesgerichtshof: *Urteil des Bundesgerichtshofs. Aktenzeichen 4 StR 287/72. Seite 383.*, 27. Juli 1972.
- [25] Catalá-Prat, A.: *RetroCamLD - A Lane Detection System Based on the Rear View Camera*. Doktorarbeit, Universitat de les Illes Balears, Palma de Mallorca, 30 Jun. 2005.
- [26] Catalá-Prat, A., F. Köster und R. Reulke: *Early Detection of Hazards in Driving Situations through Multi-Sensor Fusion*. In: *FISITA World Automotive Congress*, Bd. 2, S. 527–536, Munich, 2008.
- [27] Catalá-Prat, A., F. Köster und R. Reulke: *Sensorfusion für die Beschreibung der Fahrsituation mit Daten des Versuchsfahrzeugs ViewCar: Sensor Fusion for the Description of Driving Situations with the Data of the Test Vehicle ViewCar*. In: *Optische Technologien in der Fahrzeugtechnik*, S. 127–141. VDI-Verlag, 2008.
- [28] Catalá-Prat, A., F. Köster und R. Reulke: *Image And Laser Scanner Processing As Confident Cues For Object Detection In Driving Situations*. In: *Proc. of the ISPRS Com. V Symposium: "Image Engineering and Vision Metrology"*. 2010.
- [29] Catalá-Prat, A., F. Köster und R. Reulke: *Objektverfolgung und -fusion für die Ableitung*

- von Fahrsituationen. In: *3D-NordOst 2010, Tagungsband, 13. Anwendungsbezogener Workshop zur Erfassung, Modellierung, Verarbeitung und Auswertung von 3D-Daten*, 2010.
- [30] Catalá-Prat, A., J. Rataj und R. Reulke: *Self-Calibration System for the Orientation of a Vehicle Camera*. In: *Proc. of the ISPRS Com. V Symposium: "Image Engineering and Vision Metrology"*, Bd. 36, S. 68–73. 2006.
- [31] Cenedese, A., R. Frezza, E. Campana, G. Gennari und G. Raccanelli: *Building a Normality Space of Events: A PCA Approach to Event Detection*. In: Ranchordas, A. N. und H. J. Araújo (Hrsg.): *VISAPP 2008*, Bd. 1, S. 551. INSTICC PRESS, 2008.
- [32] Chen, Z.: *Local observability and its application to multiple measurement estimation*. IEEE Transactions on Industrial Electronics, 38(6):491–496, 1991.
- [33] Chong, C.-Y., K.-C. Chang, S. Mori und W. H. Barker: *Architectures and Algorithms for Track Association and Fusion*. Aerospace and Electronic Systems Magazine, IEEE, 15(1):5–13, 2000.
- [34] Collado, J. M., C. Hilario, J. M. Armingol und de la Escalera, Arturo: *On Board Camera Perception and Tracking of Vehicles*. In: *VISAPP (Workshop on Robot Vision)*, S. 57–66, 2007.
- [35] Cramer, H.: *Modelle zur multisensoriellen Erfassung des Fahrzeugumfeldes mit Hilfe von Schätzverfahren*, Bd. 2 d. Reihe Forschungsberichte der Professur Nachrichtentechnik. Shaker, Aachen, 2006.
- [36] Crane, C. D., D. G. Armstrong, M. Ahmed, S. Solanki, D. MacArthur, E. Zawodny, S. Gray, T. Petroff, M. Grifis und C. Evans: *Development of an integrated sensor system for obstacle detection and terrain evaluation for application to unmanned ground vehicles*. Proc. SPIE, 5804:156–165, 2005.
- [37] Dagan, E., O. Mano, G. P. Stein, A. Shashua, M. Ltd und I. Jerusalem: *Forward collision warning with a single camera*. In: *IEEE Intelligent Vehicles Symposium*, S. 37–42, Piscataway, NJ, 2004. IEEE Operations Center.
- [38] Darms, M., M. Komar und S. Lücke: *Gridbasierte Straßenrandschätzung für ein Spurhaltesystem*. In: Stiller, C. und M. Maurer (Hrsg.): *Workshop Fahrerassistenzsysteme*, Karlsruhe, 2009. Fmrt.
- [39] Dickmann, J., N. Appenrodt, O. Löhlein, M. Mekhaie, M. Mählich, M. Muntzinger, W. Ritter, R. Schweiger und S. Hahn: *Sensorfusion as key technology for future driver assistance systems*. In: *Optische Technologien in der Fahrzeugtechnik*, S. 67–84, Düsseldorf, 2008. VDI-Verlag.
- [40] Dickmanns, E. D.: *Dynamic Vision for Perception and Control of Motion*. Springer-11647 /Dig. Serial]. Springer-Verlag London Limited, London, 2007.
- [41] Dietmayer, K. C., A. Kirchner und N. Kämpchen: *Fusionsarchitekturen zur Umfeldwahrnehmung für zukünftige Fahrerassistenzsysteme*. In: Maurer, M. und C. Stiller (Hrsg.): *Fahrerassistenzsysteme mit maschineller Wahrnehmung*, S. 59–88. Springer, Berlin, 2005.

- [42] Dietmayer, K. C., J. Sparbert und D. Streller: *Model based Object Classification and Object Tracking in Traffic scenes from Range Images*. In: *IEEE Intelligent Vehicles Symposium*, S. 1–6, 2001.
- [43] Drolet, L., F. Michaud und J. Côté: *Adaptable sensor fusion using multiple Kalman filters*. In: *Intelligent Robots and Systems (IROS 2000)*, Bd. 2, S. 1434–1439, Takamatsu, 2000.
- [44] Dron, L.: *Dynamic camera self-calibration from controlled motion sequences*. In: *Computer Vision and Pattern Recognition*, S. 501–506, New York, 1993.
- [45] Effertz, J.: *Sensor Architecture and Data Fusion for Robotic Perception in Urban Environments at the 2007 DARPA Urban Challenge*. In: Sommer, G. und R. Klette (Hrsg.): *Robot vision*, Bd. 4931 d. Reihe *Lecture notes in computer science*, S. 275–290, Berlin, 2008. Springer-Verlag.
- [46] Eidehall, A.: *An automotive lane guidance system*. Doktorarbeit, Linköpings Universitet, Linköping, Sweden, 2004.
- [47] Eidehall, A.: *Tracking and threat assessment for automotive collision avoidance*. Doktorarbeit, Linköpings Universitet, Linköping, Sweden, 2007.
- [48] Eidehall, A. und L. Petersson: *Threat assessment for general road scenes using Monte Carlo sampling*. In: *IEEE Intelligent Transportation Systems*, S. 1173–1178, 2006.
- [49] Ernst, S., C. Stiller, J. Goldbeck und C. Rössig: *Camera calibration for lane and obstacle detection*. In: *IEEE Intelligent Transportation Systems*, S. 356–361, 1999.
- [50] Fardi, B., J. Dousa, G. Wanielik, B. Elias und A. Barke: *Obstacle Detection and Pedestrian Recognition Using A 3D PMD Camera*. In: *IEEE Intelligent Vehicles Symposium*, S. 437–442, Tokyo, Japan, 2006.
- [51] Floudas, N., M. Tsogas, A. Amditis und A. Polychronopoulos: *Track Level Fusion for object Recognition in Road Environments*. In: Amditis, A. (Hrsg.): *PReVENT Fusion Forum e-Journal*, Bd. 2, S. 16–23. 2008.
- [52] Form, T.: *Caroline - ein autonom fahrendes Fahrzeug im Stadtverkehr*. In: *Automation, Assistance and Embedded Real Time Platforms for Transportation, AAET*. Gesamtzentrum für Verkehr Braunschweig e.V., 2007.
- [53] Franke, U., D. Gavrilu, S. Görzig, F. Lindner, F. Paetzold und C. Wöhler: *Autonomous Driving Approaches Downtown*. *IEEE Intelligent Systems*, 13(6), 1999.
- [54] Franke, U., S. Gehrig, H. Badino und C. Rabe: *Towards Optimal Stereo Analysis of Image Sequences*. In: Sommer, G. und R. Klette (Hrsg.): *Robot vision*, Bd. 4931 d. Reihe *Lecture notes in computer science*, S. 43–58, Berlin, 2008. Springer-Verlag.
- [55] Franke, U. und I. Kutzbach: *Fast stereo based object detection for stop&go traffic*. In: *IEEE Intelligent Vehicles Symposium*, S. 339–344, 1996.
- [56] Franke, U., C. Rabe und S. Gehrig: *Kollisionsvermeidung durch raum-zeitliche Bildanalyse (Collision Avoidance based on Space-Time Image Analysis)*. it-Information Technology,



- 49(1):25–32, 2007.
- [57] Fu, Z., W. Hu und T. Tan: *Similarity Based Vehicle Trajectory Clustering and Anomaly Detection*. In: *IEEE International Conference on Image Processing*, Bd. 2, S. 2029–2032, 2005.
- [58] Fürstenberg, K. C., P. Baraud, G. Caporaletti, S. Citelli, Z. Eitan, U. Lages und C. Lavergne: *Development of a Pre-Crash Sensorial System - the CHAMELEON Project*. In: *19. VDI/VW-Gemeinschaftstagung "Fahrzeugkonzepte für das 2. Jahrhundert Automobiltechnik"*, S. 289–310. VDI-Verlag.
- [59] Fürstenberg, K. C. und K. C. Dietmayer: *Fahrzeugumfeldsensierung mit mehrzeiligen Laserscannern*. In: *Sensoren und Messsysteme*, Düsseldorf, 2004. VDI-Verlag.
- [60] Fürstenberg, K. C. und B. Roessler: *Advances in Intersection Safety*. In: Amditis, A. (Hrsg.): *PReVENT Fusion Forum e-Journal*, Bd. 2, S. 36–40. 2008.
- [61] Gacnik, J., O. Häger und M. Hannibal: *A service-oriented system architecture for the human-centered design of intelligent transportation systems*. In: *European Conference on Human Centred Design for Intelligent Transport Systems*, S. 175–181, Lyon, France, 2008.
- [62] Gerdes, J. C. und E. J. Rossetter: *A Unified Approach to Driver Assistance Systems Based on Artificial Potential Fields*. *Journal of Dynamic Systems, Measurement, and Control*, 123:431, 2001.
- [63] Gerdes, J. C., E. J. Rossetter und U. Saur: *Combining Lanekeeping and Vehicle Following with Hazard Maps*. *Vehicle System Dynamics*, 36(4):391–411, 2001.
- [64] Giachetti, A., M. Campani und V. Torre: *The use of optical flow for road navigation*. *Robotics and Automation, IEEE Transactions on*, 14(1):34–48, 1998.
- [65] Gibson, J. J. und L. E. Crooks: *A Theoretical Field-Analysis of Automobile-Driving*. *The American Journal of Psychology*, 51(3):453–471, 1938.
- [66] Gordon, B.: *Zur Bestimmung von Messunsicherheiten terrestrischer Laserscanner*. Doktorarbeit, TU Darmstadt, Darmstadt, 2008.
- [67] Graf, T., A. Catalá-Prat und H. Philipps: *Verfahren und Vorrichtung zur Fahrspurerkennung*, *Patentschrift DE 10 2008 050 456 A 1 oder G06K 9/62*.
- [68] Gruyer, D., M. Mangeas und R. Alix: *Multi-sensors fusion approach for driver assistance systems*. In: *IEEE 10th International Workshop on Robot and Human Interactive Communication*, S. 479–486, 2001.
- [69] Gruyer, D., C. Royere und V. Cherfaoui: *Credibilist multi-sensors fusion for the mapping of dynamic environment*. In: *Third International Conference on Information Fusion, FUSION 2000*, Bd. 1, S. TUC2/16 – TUC2/23, 2000.
- [70] Hall, D. L. und J. Llinas: *Handbook of Multisensor Data Fusion*. The electrical engineering and applied signal processing series. CRC Press, Boca Raton, FL., 2001.

- [71] Haralick, R. M.: *Statistical and Structural Approaches to Texture*. Proceedings of the IEEE, 67(5):786–804, 1979.
- [72] Hervieu, A., P. Bouthemy und J.-P. Le Cadre: *Video Event Classification and Detection Using 2D Trajectories*. In: Ranchordas, A. N. und H. J. Araújo (Hrsg.): *VISAPP 2008*, Bd. 2, S. 158–166. INSTICC PRESS, 2008.
- [73] Holt, V. von: *Integrale Multisensorielle Fahrumgebungserfassung nach dem 4D-Ansatz*. Doktorarbeit, Universität der Bundeswehr München, München, 2004.
- [74] Holzmann, F.: *Cooperative Adaptation between Driver and safety relevant integration of Assistant Systems as a Virtual Driver*. In: *Automation, Assistance and Embedded Real Time Platforms for Transportation, AAET*, S. 41–53. Gesamtzentrum für Verkehr Braunschweig e.V., 2007.
- [75] Jähne, B., R. Massen, B. Nickolay und H. Scharfenberg: *Technische Bildverarbeitung - Maschinelles Sehen*. Springer-Verlag, Berlin, 1996.
- [76] Jansson, J.: *Collision avoidance theory with application to automotive collision mitigation*. Doktorarbeit, Linköpings Universitet, Linköping, Sweden, Juni 2005.
- [77] Jordan, R., U. Ahlrichs, T. Leineweber, Lucas B. und P. Knoll: *Hindernisklassifikation von stationären Objekten auf Basis eines Long-Range-Radar-Sensors*. In: *Workshop Fahrerassistenzsysteme*, S. 153–161. 2006.
- [78] Kalman, R. E.: *A new approach to linear filtering and prediction problems*. Journal of Basic Engineering, 82(1):35–45, 1960.
- [79] Kämpchen, N. und K. C. Dietmayer: *Low-Level Sensorfusion von Video- und Entfernungsmessdaten zur robusten Umfelderkennung in Straßenverkehrsszenen*. In: Universität Ulm (Hrsg.): *Systemarchitekturen zur Gewährleistung sicherer und Ressourcen schonender Mobilität im Straßenverkehr*. Reischensburg, 2005.
- [80] Kämpchen, N., K. C. Fürstenberg und K. C. Dietmayer: *Ein Sensorfusionssystem für automotive Sicherheits- und Komfortapplikationen*. In: *Aktive Sicherheit durch Fahrerassistenz*, Ulm, 2004. Universität Ulm.
- [81] Kato, T., A. Kobayasi, H. Hase und M. Yoneda: *An experimental consideration for road guide sign understanding in ITS*. In: *IEEE 5th International Conference on Intelligent Transportation Systems*, S. 268–273, 2002.
- [82] Kato, T., Y. Ninomiya und I. Masaki: *An obstacle detection method by fusion of radar and motion stereo*. Intelligent Transportation Systems, IEEE Transactions on, 3(3):182–188, 2002.
- [83] Klappstein, J., F. Stein und U. Franke: *Flussbasierte Eigenbewegungsschätzung und Detektion von fremdbewegten Objekten*. In: *Workshop Fahrerassistenzsysteme*, S. 78–88. 2006.
- [84] Klaus, F.: *Einführung in Techniken und Methoden der Multisensor-Datenfusion*. Doktorarbeit, Siegen, November 1999.

- [85] Klein, L. A.: *Sensor and Data Fusion: A Tool for Information Assessment and Decision Making*, Bd. PM138. SPIE-International Society for Photo-Optical Instrumentation Engineers, 2004.
- [86] Kolski, S., M. Bellino und R. Siegwart: *Image Understanding and Data Fusion for Driver Assistant Systems*. In: *Automation, Assistance and Embedded Real Time Platforms for Transportation, AAET*, S. 28–39. Gesamtzentrum für Verkehr Braunschweig e.V., 2007.
- [87] Krüger, W.: *Robust real-time ground plane motion compensation from a moving vehicle*. *Machine Vision and Applications*, 11(4):203–212, 1999.
- [88] Labayrade, R., C. Royere und D. Aubert: *A collision mitigation system using laser scanner and stereovision fusion and its assessment*. In: *IEEE Intelligent Vehicles Symposium*, S. 441–446, 2005.
- [89] Labayrade, R., C. Royere, D. Gruyer und D. Aubert: *Cooperative Fusion for Multi-Obstacles Detection With Use of Stereovision and Laser Scanner*. *Autonomous Robots*, 19(2):117–140, 2005.
- [90] Langheim, J., A. Buchanan, U. Lages und M. Wahl: *CARSENSE - New environment sensing for advanced driver assistance systems*. In: *IEEE Intelligent Vehicles Symposium*, S. 89–94, 2001.
- [91] Liang, Y.-M., H.-R. Tyan, S.-L. Chang, H.-Y. M. Liao und S.-W. Chen: *Video stabilization for a camcorder mounted on a moving vehicle*. *IEEE Transactions on Vehicular Technology*, 53(6):1636–1648, 2004.
- [92] Lindner, P., U. Scheunert und E. Richter: *Multi Level Fusion for Environment Recognition*. In: Amditis, A. (Hrsg.): *PReVENT Fusion Forum e-Journal*, Bd. 2, S. 24–30. 2008.
- [93] Lippok, A.: *Untersuchung von Texturmerkmalen auf ihre Eignung für die Segmentierung von Bildern*. Doktorarbeit, Humboldt-Universität zu Berlin, Berlin, 2006.
- [94] Loose, H. und U. Franke: *Fahrbahnerkennung auf Landstraßen*. In: Stiller, C. und M. Maurer (Hrsg.): *Workshop Fahrerassistenzsysteme*, Karlsruhe, 2009. Fmrt.
- [95] Lopez-Orozco, J. A., de la Cruz, J.M., E. Besada und P. Ruiperez: *An Asynchronous, Robust, and Distributed Multisensor Fusion System for Mobile Robots*. *The International Journal of Robotics Research*, 19(10):914–932, 2000.
- [96] Luhmann, T.: *Nahbereichsphotogrammetrie: Grundlagen, Methoden und Anwendungen*. Wichmann, Heidelberg, 2, überarb. Aufl. Aufl., 2003.
- [97] Luong, Q. T. und O. D. Faugeras: *Self-calibration of a moving camera from point correspondences and fundamental matrices*. *International Journal of Computer Vision*, 22(3):261–289, 1997.
- [98] Mahalanobis, P. C.: *On the generalized distance in statistics*. *Proceedings of the National Institute of Science, Calcutta*, 2(1):49–55, 1936.
- [99] Mählich, M., W. Ritter und K. C. Dietmayer: *ACC Vehicle Tracking with Joint Multi-*

- sensor Multitarget Filtering of State and Existence*. In: Polychronopoulos, A. (Hrsg.): *PReVENT Fusion Forum e-Journal*. 2006.
- [100] Mählich, M., W. Ritter und K. C. Dietmayer: *Simultane Minimierung von Existenz- und Zustandsunsicherheiten in der Fahrzeugumfelderfassung*. In: *Workshop Fahrerassistenzsysteme*. 2008.
- [101] Mählich, M., R. Schweiger, W. Ritter und K. C. Dietmayer: *Sensorfusion using spatio-temporal aligned video and lidar for improved vehicle detection*. In: *IEEE Intelligent Vehicles Symposium*, S. 424–429, Tokyo, Japan, 2006.
- [102] Mählich, M., M. Szczot, M. Munz und K. C. Dietmayer: *Simultaneous Processing of Multitarget State Measurements and Object Individual Sensory Existence Evidence with the Joint Integrated Probabilistic Data Association Filter*. In: *5th International Workshop on Intelligent Transportation (WIT 2008)*, Hamburg, 2008.
- [103] Marchal, P., D. Gavrila, L. Letellier, M.-M. Meinecke, R. Morris und M. Töns: *SAVE-U: An innovative sensor platform for vulnerable road user protection*. *Intelligent Transport Systems and Services*, S. 16–20, 2003.
- [104] Marstaller, R. und H. Bubb: *Ansatz zur Bewertung von Risiko und Sicherheit beim Einsatz von Fahrerassistenzsystemen: Development of a risk analysis concept during driving with advanced driving assistance systems*. In: *Integrierte Sicherheit und Fahrerassistenzsysteme*, S. 79–97. VDI-Verlag, 2004.
- [105] Maybank, S. J. und O. D. Faugeras: *A theory of self-calibration of a moving camera*. *International Journal of Computer Vision*, 8(2):123–151, 1992.
- [106] Maybeck, P. S.: *Stochastic models, estimation, and control*, Bd. 1 d. Reihe *Mathematics in science and engineering*. Academic Press, New York, 1979.
- [107] Mendes, A., L. C. Bento und U. Nunes: *Multi-target detection and tracking with a laser scanner*. In: *IEEE Intelligent Vehicles Symposium*, S. 796–801, Piscataway, NJ, 2004. IEEE Operations Center.
- [108] Meysel, F.: *Map-Based Multi-Object Dynamic Scene Description and Atypical Event Detection*. Doktorarbeit, Humboldt-Universität zu Berlin, Berlin, 2008.
- [109] Mutambara, A. G. O.: *Design and analysis of control systems*. CRC Press, Boca Raton, FL., 1999.
- [110] Neumaier, S. und G. Färber: *Videobasierte 4D-Umfelderfassung für erweiterte Assistenzfunktionen: Vision-Based 4D-Environmental Perception for Advanced Assistance Functions*. *it-Information Technology*, 49(1):33–39, 2007.
- [111] Neumaier, S., P. Harms und G. Färber: *Videobasierte Umfelderfassung zur Fahrerassistenz*. In: *Workshop Fahrerassistenzsysteme*, S. 89–99. 2006.
- [112] Niemeier, W.: *Ausgleichsrechnung: Statistische Auswertemethoden*. de Gruyter, Berlin, 2., überarb. und erw. Aufl. Aufl., 2008.

- [113] Nieto, J., J. Guivant und E. Nebot: *The HYbrid Metric Maps (HYMMs): A Novel Map Representation for DenseSLAM*. In: *IEEE International Conference on Robotics and Automation*, Bd. 1, S. 391–396. IEEE, 2004.
- [114] Pauwels, K. und M. M. van Hulle: *Segmenting Independently Moving Objects from Ego-motion Flow Fields*. In: *Early Cognitive Vision Workshop*, 2004.
- [115] Polychronopoulos, A., U. Scheunert und F. Tango: *Centralized data fusion for obstacle and road borders tracking in a collision warning system*. In: *7th International Conference on Information Fusion*, S. 760–767, Stockholm, Sweden, 2004.
- [116] Polychronopoulos, A., M. Tsogas, A. Amditis, U. Scheunert, L. Andreone und F. Tango: *Dynamic situation and threat assessment for collision warning systems: the EUCLIDE approach*. In: *IEEE Intelligent Vehicles Symposium*, S. 636–641, Piscataway, NJ, 2004. IEEE Operations Center.
- [117] Rao, B. S., H. F. Durrant-Whyte und J. A. Sheen: *A fully decentralized multi-sensor system for tracking and surveillance*. *The International Journal of Robotics Research*, 12(1):20–44, 1993.
- [118] Rasmussen, C.: *Grouping dominant orientations for ill-structured road following*. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2004*, Bd. 1, S. I–470 – I–477. IEEE Computer Society, 2004.
- [119] Reason, J.: *Human error*. Cambridge University Press, Cambridge, 18th print Aufl., 2007.
- [120] Reichardt, D.: *Kontinuierliche Verhaltenssteuerung eines autonomen Fahrzeugs in dynamischer Umgebung*. Doktorarbeit, Universität Kaiserslautern, Kaiserslautern, 1996.
- [121] Reulke, R., F. Meysel und S. Bauer: *Situation Analysis and Atypical Event Detection with Multiple Cameras and Multi-Object Tracking*. In: Sommer, G. und R. Klette (Hrsg.): *Robot vision*, Bd. 4931 d. Reihe *Lecture notes in computer science*, S. 234–247, Berlin, 2008. Springer-Verlag.
- [122] Ringbeck, T., B. Hagebeuker, H. Kraft und M. Paintner: *PMD-basierte 3D-Optosensoren zur Fahrzeugumfelderfassung*. *Sensoren im Automobil II*, S. 209–222, 2007.
- [123] Röckl, M., J. Gacnik, J. Schomerus, T. Strang und M. Kranz: *Sensing the environment for future driver assistance combining autonomous and cooperative appliances*. In: *4th Workshop on Vehicle to Vehicle Communications (V2VCOM 2008)*, Eindhoven, Netherlands, 2008. Eindhoven University of Technology.
- [124] Röckl, M., T. Strang, M. Kranz und G. A. Center: *V2V Communications in Automotive Multi-Sensor Multi-Target Tracking*. In: *IEEE 68th Vehicular Technology Conference*, S. 1–5, Calgary, BC, 2008.
- [125] Russ, J. C.: *The Image Processing Handbook*. CRC Press, North Carolina, 2006.
- [126] Saneyoshi, K.: *3-D image recognition system by means of stereoscopy combined with ordinary image processing*. In: *IEEE Intelligent Vehicles Symposium*, S. 13–18. IEEE, 1994.

- [127] Sattel, T. und T. Brandt: *Ein Beitrag zur Interaktiven Querfuhrungsassistenz von Kraftfahrzeugen*. In: *Integrierte Sicherheit und Fahrerassistenzsysteme*, S. 207–222. VDI-Verlag, Düsseldorf, 2006.
- [128] Schiementz, M.: *Postprocessing architecture for an automotive radar network*. Doktorarbeit, Technische Universität Hamburg-Harburg, Hamburg-Harburg, 2005.
- [129] Schulze, M., G. Nocker und K. Bohm: *PReVENT: A European program to improve active safety*. In: *5th International Conference on Intelligent Transportation Systems Telecommunications*, Brest, France, 2005. PReVENT.
- [130] Schwarzing, M.: *Flexible Fahrzeugerkennung für das Rechnersehen in einem Automobil*, Bd. 375 d. Reihe *Fortschritt-Berichte VDIReihe 10, Informatik/Kommunikationstechnik*. VDI-Verlag, Düsseldorf, 1995.
- [131] Schweiger, R., H. Neumann und W. Ritter: *Multiple-cue data fusion with particle filters for vehicle detection in night view automotive applications*. In: *IEEE Intelligent Vehicles Symposium*, S. 753–758, 2005.
- [132] Shimomura, N., K. Fujimoto, T. Oki und H. Muro: *An Algorithm for Distinguishing the Types of Objects on the Road Using Laser Radar and Vision*. *IEEE Transactions on Intelligent Transportation Systems*, 3(3):189–195, 2002.
- [133] Simon, A. und J.-C. Becker: *Vehicle Guidance for an Autonomous Vehicle*. In: *IEEE Intelligent Transportation Systems*, 1999.
- [134] Skiba, R.: *Die Gefahrenträgertheorie*, Bd. 106 d. Reihe *Forschungsbericht (Bundesanstalt für Arbeitsschutz und Unfallforschung)*. 1973.
- [135] Skutek, M.: *Ein PreCrash-System auf Basis multisensorieller Umgebungserfassung*. Doktorarbeit, Technischen Universität Chemnitz, Chemnitz, 2006.
- [136] Skutek, M. und D. Linzmeier: *Fusion von Sensordaten am Beispiel von Sicherheitsanwendungen in der Automobiltechnik*. *at - Automatisierungstechnik*, 53(7):295–305, 2005.
- [137] Steffel, H.-C., D. Wagner und D. Wagner: *Danger Recognition System for Vehicles, comprising at least one lateral and rear environment detection unit*. WO Patent WO/2004/068, 2004.
- [138] Steinbrecher, R.: *Bildverarbeitung in der Praxis*. R. Oldenbourg Verlag, Oldenbourg, 1993.
- [139] Stiller, C.: *Fahrerassistenzsysteme - Von realisierten Funktionen zum vernetzt wahrnehmenden, selbstorganisierenden Verkehr*. In: Maurer, M. und C. Stiller (Hrsg.): *Fahrerassistenzsysteme mit maschineller Wahrnehmung*, S. 1–20. Springer, Berlin, 2005.
- [140] Stiller, C., J. Hipp, C. Rössig und A. Ewald: *Multisensor obstacle detection and tracking*. *Image and Vision Computing*, 18(5):389–396, 2000.
- [141] Stiller, C., S. Kammel, I. Lulcheva und J. Ziegler: *Probabilistische Methoden in der Umfeldwahrnehmung Kognitiver Automobile Probabilistic Methods for Environment Perception of Cognitive Automobiles*. *at-Automatisierungstechnik*, 56(11):563–574, 2008.

- [142] Stiller, C., S. Kammel, B. Pitzer, J. Ziegler, M. Werling, T. Gindele und D. Jagszent: *Team AnnieWAY's Autonomous System*. In: Sommer, G. und R. Klette (Hrsg.): *Robot vision*, Bd. 4931 d. Reihe *Lecture notes in computer science*, S. 248–259, Berlin, 2008. Springer-Verlag.
- [143] Streller, D. und K. C. Dietmayer: *Lidar-Sensoren zur Umfelderkennung bei Straßenfahrzeugen*. In: *Messtechnisches Symposium des AHMT*, S. 163–176, Kassel, 2002.
- [144] Stüker, D.: *Heterogene Sensordatenfusion zur robusten Objektverfolgung im automobilen Straßenverkehr*. Doktorarbeit, Universität Oldenburg, Oldenburg, 2004.
- [145] Thomanek, F.: *Visuelle Erkennung und Zustandsschätzung von mehreren Straßenfahrzeugen zur autonomen Fahrzeugführung: Dissertation*, Bd. 272 d. Reihe *Fortschritt-Berichte VDIReihe 12, Verkehrstechnik/Fahrzeugtechnik*. VDI-Verlag, 1996.
- [146] Tomasi, C. und T. Kanade: *Detection and tracking of point features*. School Comput. Sci., Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep. CMU-CS-91-132, 1991.
- [147] Ulrich, I. und I. Nourbakhsh: *Appearance-based obstacle detection with monocular color vision*. In: *17th National Conference on Artificial Intelligence (AAAI-2000) / 12th Conference on Innovative Applications of Artificial Intelligence (IAAI-2000)*, S. 866–871, 2000.
- [148] Utasi, A. und L. Czúni: *Anomaly Detection with Low-Level Processes in Videos*. In: Ranchordas, A. N. und H. J. Araújo (Hrsg.): *VISAPP 2008*, Bd. 2, S. 678–681. INSTICC PRESS, 2008.
- [149] VDE, DIN: *Allgemeine Leitsätze für das sicherheitsgerechte Gestalten technischer Erzeugnisse – Teil 2: Begriffe der Sicherheitstechnik–Grundbegriffe*, 1987.
- [150] Ventures, M. D.: *Stanley: The Robot that Won the DARPA Grand Challenge*. *Journal of Field Robotics*, 23(9):661–692, 2006.
- [151] Vercauteren, T. und X. Wang: *Decentralized sigma-point information filters for target tracking in collaborative sensor networks*. *IEEE Transactions on Signal Processing*, 53(8 Part 2):2997–3009, 2005.
- [152] Viola, P. und M. Jones: *Rapid object detection using a boosted cascade of simple features*. In: *IEEE Conference on Computer Vision and Pattern Recognition*, S. 511–518, 2001.
- [153] Vogel, K.: *A comparison of headway and time to collision as safety indicators*. *Accident Analysis and Preventions*, 35(3):427–433, 2003.
- [154] Vu, T.-D., O. Aycard und N. Appenrodt: *Online Localization and Mapping with Moving Object Tracking in Dynamic Outdoor Environments*. In: Amditis, A. (Hrsg.): *PRe-VENT Fusion Forum e-Journal*, Bd. 2, S. 3–8. 2008.
- [155] Walchshäusl, L., R. Lindl, K. Vogel und T. Tatschke: *Detection of Road Users in Fused Sensor Data Streams for Collision Mitigation*. In: *Advanced Microsystems for Automotive Applications 2006*, Bd. 2 d. Reihe *VDI-Buch*, S. 53–65. Springer Berlin Heidelberg, 2006.

- [156] Weiss, T. und K. C. Dietmayer: *Positionierung eines Fahrzeugs in unbekannten Gebieten mit Hilfe von Laserscannern Localization in Unknown Environments Using Laser Scanner*. at-Automatisierungstechnik, 56(11):554–562, 2008.
- [157] Weiss, T., J. Huchler und K. C. Dietmayer: *Generierung präziser digitaler Verkehrsinfrastrukturkarten mit Laserscannern*. In: *Integrierte Sicherheit und Fahrerassistenzsysteme*, S. 239–257. VDI-Verlag, Düsseldorf, 2006.
- [158] Weiss, T., S. Wender, B. Schiele, K. C. Fürstenberg und K. C. Dietmayer: *Online Mapping for the Robust Detection of Driving Corridors*. In: Amditis, A. (Hrsg.): *PReVENT Fusion Forum e-Journal*, Bd. 2, S. 41–47. 2008.
- [159] Wender, S. und K. C. Dietmayer: *3D Vehicle Detection Using a Laser Scanner and a Video Camera*. In: *6th European Congress on Intelligent Transport Systems*, 2007.
- [160] Wender, S. und K. C. Dietmayer: *3D vehicle detection using a laser scanner and a video camera*. *Intelligent Transportation Systems*, 2(2):105–112, 2008.
- [161] Wender, S., K. C. Fürstenberg und K. C. Dietmayer: *Object Tracking and Classification for Intersection Scenarios Using A Multilayer Laserscanner*. In: *11th World Congress on Intelligent Transportation Systems (ITS 2004)*, Nagoya, Japan, 2004.
- [162] Wender, S., T. Weiss, K. C. Fürstenberg und K. C. Dietmayer: *Object Classification exploiting High Level Maps of Intersections*. In: *Advanced Microsystems for Automotive Applications 2006*, Bd. 2 d. Reihe VDI-Buch, S. 187–203. Springer Berlin Heidelberg, 2006.
- [163] Wetzel, D.: *Wissensbasierte Verkehrsszenenanalyse zur Fahrerunterstützung*. Doktorarbeit, Universität Erlangen-Nürnberg, Sankt Augustin, 1995.
- [164] Wisselmann, D., K. Gresser, H. Spannheimer, K. Bengler und A. Huesmann: *Connected-Drive - ein methodischer Ansatz für die Entwicklung zukünftiger Fahrerassistenzsysteme*. In: *Tagung Aktive Sicherheit durch Fahrerassistenz*, München, 2004.
- [165] Wohlfeil, J.: *Erkennung und Verfolgung von Fahrzeugen mit bewegter Kamera*. Doktorarbeit, Humboldt-Universität zu Berlin, Berlin, 2006.
- [166] Wohlfeil, J. und R. Reulke: *Erkennung und Verfolgung von Fahrzeugen mit mobiler Kamera: Detection and Tracking of Vehicles with a Moving Camera*. In: *Optische Technologien in der Fahrzeugtechnik*, S. 181–195. VDI-Verlag, 2008.
- [167] Yan, L. P., B. S. Liu und D. H. Zhou: *The modeling and estimation of asynchronous multirate multisensor dynamic systems*. *Aerospace Science and Technology*, 10(1):63–71, 2006.
- [168] Yang, J. H.: *Development of a performance-based approach for collision avoidance and mitigation*. Doktorarbeit, 2003.
- [169] Yang, L., J. H. Yang, E. Feron und V. Kulkarni: *Development of a performance-based approach for a rear-end collision warning and avoidance system for automobiles*. In: *IEEE Intelligent Vehicles Symposium*, S. 316–321, Columbus, OH, 2003.